

Good news for the year ahead

The traditional sport of guessing what will happen in the next twelve months is more cheerful this year than in the recent past. But anything could go wrong.

John Maddy

THE best augury for 1987 is the return of Andreii Sakharov to Moscow. The next best is that Sakharov himself, while complaining that there are many others like him still kept against their inclinations in places even more uncomfortable than Gor'kii, should have praised the changes in Soviet society brought about, he says, by the new general secretary, Mr Mikhail Gorbachev. The tribute is all the more remarkable because Sakharov, a card-carrying truth-teller for four years isolated in a militarily-sensitive provincial city, could not within ten minutes of stumbling from an overnight train have founded his approval on personal observation. In any case, as Sakharov will soon discover, not much has changed as yet. The tyranny of the bureaucracy, the stultifying successor of the dictatorship of the proletariat, is still largely intact. What Sakharov is saying is no more than that he now shares with a small but influential band of his compatriots the belief that there is hope. Is that not grand?

Of course. But is it necessarily good? There are many outside the Soviet Union who consider Gorbachev's arrival to have been a misfortune; having learned how to handle septuagenarian emphysemiacs, people do not like seeing the goal posts moved. Yet the potential benefits of change far outweigh the inconvenience. The Soviet Union is a modern state with an antiquated economy which, some years, cannot feed even itself; who, elsewhere, could genuinely cry foul if it could be shamed into assisting the cause of development as other countries must, with cash and not simply with automatic rifles? The Soviet Union, forever talking world peace, repeatedly embarks on breathtakingly foolhardy military projects such as that, a decade ago and while the SALT II treaty was still fresh, to aim more than 200 SS-20 missiles at Western Europe; has the unrequited unilaterally-Soviet ban on nuclear tests for the past 18 months been an attempt to show that somebody values butter more than guns?

History will no doubt explain, but there are more immediate benefits to be won. The virtual isolation for more than half a century of the intellectual community of the Soviet Union from that of the rest of the world has been a great disservice to both parts; will 1987 see a significant erosion of that artificial barrier? The prospects look good. And will the tendency to openness (see p. 9) persist? Sakharov's return to Moscow is a pointer in the right direction; evidently the Soviet authorities have reconciled themselves to the notion that there should be an open critic let loose to talk to whom he chooses, just as governments elsewhere have to put up with endless criticism — the Strategic Defense Initiative will not work, secretly selling arms to your public enemies is mistaken and so on. If it carries on like this, the Soviet Union may one day risk a free press. Meanwhile, one consequence cannot but be a beneficial access of civility. Further ahead is the prospect of the mutual stimulus that will come from allowing two separate intellectual communities to rub shoulders with each other.

Realism

Elsewhere, the auguries are also not so bad. Realism seems to be back in fashion. The lofty detachment with which the US administration has dealt with the Congress for the past six years will

be transformed, now that the Democrats control the Senate, into a process by which the two sides reach an accommodation by familiar bargaining. The new Congress could well be able to make the point that a workable agreement on strategic arms would be preferable to the preservation of planned spending on the Strategic Defense Initiative; the case will be easier to make now that the administration has been convicted of naivety in its dealings with Iran. The other side of this coin is that the Congress will be in no mood to control its spending in the two years preceding the next presidential election, so that a substantially shrunken federal deficit is a poor prospect; that matters because the present state of affairs in which the United States is sustained by borrowing from abroad cannot last much longer. The new Congress will be more eager than its predecessor, and probably more successful, at enacting short-term protectionist palliatives but, luckily, it seems now to have sunk home that righting the economic imbalance between the United States and the rest of the world requires reform of the educational system at all levels. That cannot but be good for scholarship and research.

Problem

Japan, the chief financier of the US trade deficit, is the image in this looking-glass. Superficially, the problem for Japan is how to spend all the wealth its companies have earned in the world's export markets. Individuals will not change their ways, but will carry on saving. Companies, fearful that their export trade may be grinding to a halt, are busily and ingeniously tunnelling beneath the now-high yen, building factories overseas. The year ahead could be that in which people elsewhere who now acknowledge that they owe their video-recorders to Japan must also learn to thank Japanese companies for their jobs. The absorbing question is whether one outcome will be the long overdue reform of the Japanese universities. Another is whether Japan will realize its ambition of making a decisive contribution to basic science. It is a sign of grace that these have become political goals.

Europe (see p. 8) is less confident. The management of the European Communities repeatedly perplexes those who grapple with the task. But national politicians have made a useful start on dealing with the recurring agricultural surpluses. They are also beginning to realize what they should have known 30 years ago, that there is no point in belonging to a common market if public purchases are still made chauvinistically; there should be some progress on that front this year. The European Commission's frustration that its ambitions for Community research have not been denied but simply ignored, by the cancellation of last month's planned meeting of research ministers, is understandable but beside the point. It is more important that some member states, even poor Britain for example (see p. 5), have come round to believing that the way forward entails the more deliberate harnessing of innovation to survival, economic and otherwise. These days, after all, it is a luxury to believe that a recipe that may work. Would that there were such a one for, say Africa. Japan's New Year loan of \$9,000 million to this international aid agencies will help, but only as a palliative. Self-sustainment is the crying need. □



Selling off the silver

The fashion for selling public assets is spreading, but governments need rules for doing so.

BRITAIN, which used to boast of being a source of innovation quickly exploited by others, may now at least take a little pride at the way in which its government's schemes for selling nationalized industries to private individuals and corporations is being widely copied. Seven years ago, at the beginning of the Thatcher administration, the new government made a careful hunt through the national accounts for ways of reducing public expenditure. The motive was not so much ideological, to get government off the people's back, as practical, to reduce the proportion of the national income being spent by central government. In the event, the search was not conspicuously successful. The welfare state, the government discovered, had come to stay; even its own supporters believed that state-run schools and hospitals had become necessities.

Almost as an afterthought, attention turned to the selling of public assets, tangible (houses owned by local authorities or land) and less so (nationalized industries). Amersham International, the isotope company, was one of the first to be sold (at a price reflecting roughly half its value). Since then, the telecommunications monopoly (British Telecom) and the gas industry (British Gas) have been 'privatized'. An airline (British Airways) and an aero-engine company (Rolls-Royce) will be next.

The immediate benefits have been considerable; in the present financial year, the British government is likely to have to borrow between £2,000 and £3,000 million less than seemed likely nine months ago, and will accordingly be better placed to write a more enticing budget for the next financial year. Is it any wonder that the formula has been taken up enthusiastically elsewhere? The Japanese government has quietly privatized its telecommunications authority NTT, and is now embarking on the much more difficult task of selling off its railroads (which are unprofitable when account is taken of the cost of building all those bullet trains and the tracks on which they run). Even in the United States, the Congress is facing up to the problem of selling Conrail, the railway company left anomalously on its hands when the privately owned railways were driven to the wall 15 years ago, but may yet shrink from the dearth of buyers.

Before the fashion for selling public assets spreads too far, potential imitators of the British practice should prudently consider the darker side of this new coinage. First, there is a doctrinal difficulty. Strictly speaking, the proceeds of the sale of an asset should be regarded as a capital receipt and not a contribution to current revenue, so that the £8,000-million-plus that will eventually be raised by selling British Gas should be used to liquidate some of the British government's long-term debt, not to reduce income tax. But the argument is needlessly strict; the assets were originally acquired (as municipal enterprises) by the issue of government stock (which is a way of printing money) and have been augmented by public cash subsidies and the surpluses engendered by asking consumers to pay more for gas than would have been necessary to keep a privately owned enterprise afloat. Will the economic benefits of lower taxes sustain the illusions taxpayers are being offered when the backlog of assets for sale is exhausted? Few seem to ask this question.

Other difficulties stem from the nature of the assets being sold, some (but not all) of which have social or nationally strategic functions. The days have gone when national pride required that every government should own an airline (although some governments are slow to learn), but a government that sells off a telecommunications or railway network which then fails is likely to find itself compelled politically to buy it back, or to recreate it, most probably at a loss. At the present rate of the disposal of public assets, the first test-case cannot be far away. Most probably, in Britain, it will arise when some local authority considers that its people need a bus service and decides to

provide one partly out of public funds. That would bring the wheel full circle; once municipally-owned bus companies, nationalized by a previous government, are now being sold off piecemeal by the National Bus Company, which cannot know that all will survive. But the difficulty, for a government, would be more acute if a major enterprise such as a telecommunications network were to land itself in trouble.

These are general issues. The more immediate issue is the degree to which newly privatized enterprises should be free to milk their markets for what they can. Privatized British Telecom is forbidden to raise its charges by more than three per cent less than the going rate of inflation in its first five years, is regulated in its administrative decisions by an effective public commission and has an embryo competitor (called Mercury) breathing down its neck. Other arrangements, such as those for British Gas, are less effective safeguards against monopoly power. Those chickens may yet come home to roost.

Circumnavigation again

The successful flight around the Earth of the Voyager aircraft deserves even more attention.

WHO could wish to spend nine days on the circumnavigation of the Earth in the uncomfortable cabin of a lightweight propeller-driven aircraft, technically intermediate between a conventional aircraft and a motorized kite? These days, there would be little difficulty in making the trip by scheduled commercial flights in less than two days. People who travel in Earth satellites take little longer than ninety minutes on the journey. The two-person crew of Voyager, which returned to Edwards Air Force Base in California on 23 December, seem to have set great store on the circumstance that they were able to make the journey without refuelling their aircraft, but even that is unremarkable these days, when the refuelling of military aircraft in flight happens regularly. So what is remarkable about Voyager?

It is touching that much of the United States seems to have regarded the flight as the successful culmination of a national endeavour; a presumably hard-boiled spokesman of the Federa-



Pilots Dick Rutan and Jeana Yeager celebrate.

tion of American Scientists is even quoted as believing that the aircraft would partly make up for Challenger, the shuttle that went up in flames nearly two years ago. The truth is that the went up in flames. The truth is that the flight was indeed a considerable adventure, and none the worse for that. It was also a compelling demonstration of a truth too often overlooked, that there is still almost literally mileage in the development of conventional aircraft, not least by the use of composite materials to reduce their weight and by the exploitation of novel aerodynamics. The cleverness of the Voyager design lies in the hugeness of the lifting area per unit weight. But what the world needs is an aircraft capable of using these techniques (and others) for circumnavigation in as many hours as the days spent by Voyager. □

Lord Stockton, president of Macmillan publishers, founders and owners of *Nature*, died at the age of 92 early this week. As Harold Macmillan he was Britain's prime minister from 1959 to 1963. □

Sakharov — a wind of change?

- Sakharov's internal exile ends at last
- Campaign on human rights continues
- Plans for a return to science

£60

London

DR Andrei Dmitrievich Sakharov, 'father of the Soviet hydrogen bomb', Nobel Peace Prize winner and doyen of the Soviet human rights movement, is alive, reasonably well, extremely vocal, and once more living in Moscow with his wife, Yelena Bonner. His 'administrative exile', announced on 22 January 1980, has been rescinded after almost seven years, and his wife's sentence quashed so that they could return, together, to Moscow. The surprise announcement, coming at the end of a year that also saw the depar-



Sakharov has been the centre of much attention from the Western media.

ture from the Soviet Union of such notable human rights campaigners as Anatolii (now Nathan) Shcharanskii and Yuri Orlov, is seen as a major signal of a new Soviet attitude on human rights issues.

These events, however, have to be viewed against a wider background. So far, those who have benefited most have been the big names among the Soviet scientific dissidents and refuseniks, for whom major campaigns have been launched. During the past seven years, Sakharov's exile, in particular, has on several occasions been a cause of embarrassment to Soviet scientists and diplomats in their encounters with Westerners.

Scarcely a month after Sakharov left Moscow, the Hamburg Scientific Forum (part of the Helsinki review process) brought about some sharp exchanges, in which, at one point, the interpreters switched off their headphones and refused to work until the Soviet delegate had cooled down. On the literary front, the Soviet journal that had contracted, sight unseen, to publish Arthur C. Clarke's science-fiction novel *2010*, was considerably embarrassed to find that Clarke portrayed Sakharov in his ninetieth year as president of the Soviet Academy of Sciences and a pillar of the Soviet scientific establishment, a prediction that now seems considerably less fantastic than a few weeks ago.

The Soviet authorities will undoubtedly gain from his return. Not only does it reinforce the new, 'liberal' image of the Gorbachev regime, but Sakharov's views on President Reagan's Strategic Defense Initiative (SDI) in many ways coincide with the official Soviet stance. As early as 1983, before the US President announced the programme, Sakharov expressed his doubts on whether such weapons could work. Now, although he admits that such a system could be implemented, he has publicly stated that a "potential enemy with highly developed technology" could always find a means to counter such a space defence — at a considerably lower cost than the SDI system itself.

Soviet disarmament negotiators would have been less happy, however, with Sakharov's comment that it is "ungrounded" to demand that the Americans stop all research on SDI, and that it is "illegal" to demand the halting of such research as a precondition of all other arms-control negotiations. Indeed, "we may infer from general knowledge", he said, that similar research is going on in the Soviet Union, and it would be "simply unrealistic" to stop such research.

In the future, Sakharov says, he will

Robert Wainwright Third-stage management

ARIANE, the European space launch vehicle, will not be launched in February 1987 as planned. The European Space Agency (ESA) has postponed the launch of an improved Ariane with a more powerful third-stage igniter for 6 or 7 weeks while the manufacturer, Arianespace, carries out further tests. Two of the last three launches have failed due to failures of the third stage to ignite. ESA hopes to announce a new launch date later this month.

Vera Rich Supek denies charges

PROFESSOR Ivan Supek, the Yugoslav physicist and human rights campaigner, is the victim of a virulent campaign in the Yugoslav media. The attacks were triggered by an attempt by the South Slav Academy of Sciences and Arts to obtain for Supek, who retired last year at 70, a supplementary pension commensurate with his intellectual achievements. But the media, the party organ and the veterans' league SUBNOR claim that such a pension would not be appropriate. Supek, they say, while rector of Zagreb University in 1968–72, inspired his students with anti-party policies. In a telephone interview with

continue to campaign against abuses of human rights, and, in particular, for the liberation of prisoners of conscience. He also hopes to return to scientific work. This latter decision has the approval of the authorities; during the 1970s, a frequent charge levelled against Sakharov by establishment Soviet scientists was that he had "abandoned science" for human rights, while in 1980, the decision to exile him to Gor'kii was officially explained as being for his own good, to remove him from distractions such as Western journalists, so that he could continue his scientific career uninterrupted. But an active scientific career implies the right to attend scientific conferences at home and abroad, and Sakharov has already expressed a desire to travel to such conferences, as well as to visit his relatives in the United States. But, only a few months ago, an official spokesperson stated that Sakharov would never be allowed to go abroad, as he has had access to classified information.

It is understandable (although contrary to the International Covenant on Human and Civil Rights) that the Soviet authorities might not want to lose Sakharov's exceptional talents to another country, but Sakharov, it must be stressed, would prefer, all else being equal, to live in Russia. It is "trips with a return" that interest him, not emigration. And he has had no access to classified information since 1968.

Vera Rich

Nature he vehemently denied the charge. It was a time of widespread demands for greater democracy in Yugoslavia, he said but few students were directly concerned, and he himself has always been "very moderate".

Joseph Palaca Creationists back in court

THE contention that 'creation science' is religious doctrine dressed in scientific garb is among those facing US Supreme Court justices in the case of *Edwards versus Aguillard*. At issue is a Louisiana statute requiring evolution and "creation science" to be given "balanced treatment" in schools. The US constitution does not permit laws that can be seen as the "establishment of religion". The Louisiana argument is that the law is entirely secular. Judgement is expected in July, but that is unlikely to be the end of the matter.

Polish fallout corrected

IN the news item "Polish fallout underestimated" *Nature* 324, 603; 1986) the fifth sentence should begin "The emergency network activated on 29 April", and not (as printed) "The emergency on 29 April".

US-Soviet collaboration on stratospheric ozone depletion

Washington

DECLARING that there is "special value" in cooperation on global problems related to stratospheric ozone depletion, the US-USSR Joint Committee on environmental protection announced a scientific meeting on that topic to take place shortly in Moscow. Ozone investigations form one of 38 new scientific and technical projects adopted by the two countries during a meeting in Washington last month.

The Soviet Union has an extensive network of ground stations for measuring ozone, but the Soviet measurements are not made with the Dobson spectrophotometer which is standard in the West. One goal of the Moscow meeting will be to develop calibration protocols so that data can be compared. The United States will also be supplying Soviet scientists with ozonesondes for vertical ozone measurements at Soviet Antarctic stations.

In the summer, East German and US ozonesondes will be tested together in Rylsk, outside Moscow. In a related effort, the National Oceanographic and Atmospheric Administration (NOAA) will be sending a chlorofluorocarbon monitoring instrument to Estonia in the next few weeks. John Miller, deputy director of NOAA's air resources laboratory, says US scientists will have access to all data collected in Estonia. As well as the ozone studies, the joint committee also approved studies of greenhouse gases and

their effect on climate. Following a meeting in Leningrad last July, the committee concluded that human influences probably caused the 0.5 °C rise in temperature of the Northern Hemisphere observed over the past 100 years. Further changes in the balance of atmospheric gases could have a significant effect on agriculture and the coastal environment.

In late spring, a Soviet delegation will visit the United States to discuss air pollution monitoring and modelling. At about the same time, another Soviet delegation will arrive to discuss stationary sources of pollution, and a US group may visit the Soviet Union to hold talks on mobile sources of pollution.

The joint committee on cooperation in environmental protection came into being as a result of a bilateral agreement signed in 1972, and is renewable every five years. Although it came as no surprise, the two countries agreed to another five-year extension beginning in May. At present, relations are particularly cordial on environmental issues, with the Soviet Union particularly keen.

The Soviet Union has also recently announced plans to rejoin the Ocean Drilling Project, which brings to eight the number of fully fledged participants. The European Science Foundation joined last spring (see *Nature* 321, 102; 1986). An official signing ceremony is planned for Moscow later this winter. **Joseph Palca**

Pentagon move slows space station progress

Washington

THE US Department of Defense has thrown a potentially damaging wrench into the delicate machinery of international cooperation on the space station. At the Pentagon's bidding, NASA (National Aeronautics and Space Administration) has postponed a scheduled January meeting in Paris with the European Space Agency (ESA) until an interagency group can meet to determine what part the Defense Department will play in the international project.

The Pentagon asked for the delay during December. An interagency group, chaired by the Department of State, has already held one meeting to discuss the Defense Department's interest and to see how it will alter the US negotiating position. The working group will report to Air Force Colonel Gerald May, director of space for the National Security Council.

The Pentagon's interest in the space station appears to have blown up suddenly. During the past year, there has not been a strong interagency policy on the space station. But once the Pentagon indicated its concern, other agencies chimed in with other problems, although defence issues will probably dominate the working group's agenda.

Pentagon spokesman Lt Col. Donald Brownlee says no particular uses for the space station have been established. Brownlee says the department is merely making sure that its options are not foreclosed. After study, Brownlee contends, the space station may turn out to be ideal for conducting experiment for the Strategic Defense Initiative (SDI). There are no plans to use the space station as an operational base.

NASA has never concealed the possibility of the Defense Department becoming a user of the space station, but has never emphasized it. The level of military participation could cause problems for ESA and Japan. Ian Pryke, ESA's Washington representative, says the ESA covenant allows the agency to participate only in projects with peaceful purposes. ESA is already embroiled in difficult negotiations with NASA over its financial contributions to the space station (see *Nature* 324, 505; 1986). Depending on what role the Pentagon wants to play, the debate within ESA may shift to the interpretation of the term 'peaceful purposes'.

Japan has always regarded the space station as being reserved for peaceful activities. A heavy military component in the project could pose problems for Japan's continued participation. **Joseph Palca**



THE south side of Lake Nyos, Cameroon, where on 21 August 1986 a huge volume of toxic gas was released from beneath the lake, producing an aerosol of water mixed with gas that swept down the valleys to the north of the lake killing more than 1,700 people. This picture was taken by Dr S.J. Freeth on 8 September 1986. Wave damage to the vegetation, particularly on the delta to the south of the lake and over the promontory to the west, is clearly visible. Drs Freeth and R.L.F. Kay, who were in the UK team that investigated the disaster, report their conclusions in a *News and Views* article in next week's *Nature*. □

Testing-time for British industry both at home and abroad

THE year ahead in Britain will be memorable for the energy the government will spend exhorting British industry to mend its ways. Britain must think internationally, its research effort must be directed towards more commercial goals and its development budget must be substantially increased — by industry, not the taxpayer. That will be the message, based on lessons bitterly learned last year: British innovation, conspicuously less successful than that in Japan and the United States, is now also compared unfavourably with that elsewhere in Europe. There will also probably be a general election.

The lessons learned by Britain's General Electric Company (GEC) in the last two weeks of 1986 starkly typify the problem. The Ministry of Defence has been in the market for an airborne early warning radar system since 1978, since when GEC has spent £900 million on the development of a system called Nimrod. But in the event, on the advice of the Royal Air Force and other technical advisers, the government has chosen to buy the Boeing AWACS system instead.

This shaming development has confirmed the government in its view that British industry is too parochial, that its attempts to exploit research are sluggish and that research efforts in university and commercial laboratories are not properly coordinated. GEC's painful lesson will be repeated again and again this year at the hands of a government convinced that it spends too much on research and has too little to show for it.

To steer Britain in the right direction, the government plans to create a completely new mechanism, the first element of which will involve a team of industrialists, academics, financiers and bureaucrats charged with hammering out a new policy on industry and research. The foundations of this structure, unlike any previously attempted in Britain, will be laid by the spring by its chief architects, the team of scientific advisers at the Cabinet Office under the government's chief scientific adviser John Fairclough.

By the end of the year, there will also be new procedures to ensure that the funds spent through the Advisory Board for the Research Councils have the effect of concentrating research more effectively on selected centres. The Cabinet Officer believes that this strategy will ensure the existence of well-equipped centres of excellence capable of competing internationally.

Research and development in information technology may well benefit specifically this year. Much will depend on the government's response to the demand of

the IT86 Committee, under Sir Austin Bide (retired chairman of Glaxo) for an investment of £1,000 million over five years, half of it from public funds. Bide is bullish, predicting a positive response. His plan is for a continuation of the five-year Alvey programme of research in information technology, perhaps incorporating the British contribution to the European Commission's ESPRIT initiative.

But 1987 will also be a year in which Britain discovers that internationalism does not come cheap. Officials of the British National Space Centre (BNSC) are still waiting for the government's approval of their 15-year plan, which would entail an annual commitment of £200 million, most of it for collaborative projects through the European Space Agency (ESA). While the government pontificates about the insularity of British research and the lack of an international vision in industry, it has itself been rather slow to respond to this latest international opportunity.

In another field, the government's international pretensions will be tested as

Prospects for 1987

United Kingdom

Bill Johnstone in London

the European states prepare to launch the first Direct Broadcasting Satellites (DBS), Earth satellites broadcasting television signals strong enough to be picked up by individual homes equipped with small antennas. The Independent Broadcasting Authority (the licensing authority for commercial television and radio) has at last awarded a franchise to operate three DBS channels. That dubious prize, which could cost £700 million over 15 years, has gone to British Satellite Broadcasting (BSB), a consortium of Amstrad Consumer Electronics (a prominent hi-fi and microcomputer supplier), the Anglia and Granada commercial television stations, the Pearson Group (owners of the *Financial Times* among other things) and the Virgin Group (the record company that recently founded an airline).

There are potential disputes ahead over European programming policy. The European Commission's proposal that DBS should carry a specified minimum proportion of material of European origin has been contested by European states. It remains to be seen whether what will emerge from a meeting of the Council of Europe last December will be more acceptable, and whether any agreement can be policed.

Meanwhile, both France and West

Germany are to launch in the next 12 to 18 months their own DBS services, assisted by generous grants from their governments. It will be possible to receive these services in the British Isles long before the British system, due for launch in 1989–90, is in service.

Another contentious issue in Britain will be that of who builds the satellite for the BSB consortium. Five years ago, the British Broadcasting Corporation (BBC) together with a group of television and industrial companies, won a licence to op-



erate a DBS system, arranged to buy three satellites from a consortium consisting of British Aerospace, GEC-Marconi and British Telecom (the denationalized telecommunications company) and then aborted the project on the grounds of cost. Now, Comsat and RCA will be prominent bidders for the satellites that BSB will need. British industry, not for the first time, may find itself bested by the United States.

During 1987, Britain may also finally decide its long-term relationship with the European Organisation for Nuclear Research (CERN), the high-energy physics laboratory at Geneva.

Britain last year declared its wish to reduce its contribution to CERN (which reached £37.56 million in 1985–86) by 25 per cent before 1991, for which reason an internal review committee has been established to work out whether the British wish can be accommodated. But the declining value of sterling has since made membership even more expensive; virtually all the extra £20 million for research won from the government towards the end of last year will need to go to CERN. Action of some kind, perhaps drastic, is likely once the internal review is complete.

On CERN as in other matters, the government will have the final say. Will it throw in its hand on CERN for the sake of spending a little on space? Or, in an election year, will it find that there is enough to do both? A little time should provide the answer. □

Health spending still increasing but space science stalled

WHAT the US government does for (and to) science in 1987 is bound to be influenced by the imbroglio of the Iranian arms affair — the most serious political scandal for a decade, variously known as "Iran-gate", "contragate" and "Iranamok".

At the very least, the long process of inquiry will significantly affect the capacity of the President to govern. But as the fiasco has acquired its present importance because the administration chose not to tell the Congress what was going on, distrust between the two parts of government has been magnified. With both houses of Congress now controlled by Democrats, a Republican president with two years left in office will find it harder than ever to persuade congressmen to his way of thinking.

Arms control will be an early issue between these two parts of the government. More than half of the Senate's 100 members have already signed a letter urging the President to bring back the United States within the limits of the SALT II treaty; they have in mind the deployment in November of a B52 bomber equipped with nuclear-armed cruise missiles which put the United States beyond bounds. But there will be more general dissent from the administration's policies on arms control. Enthusiasm for military research and particularly for the Strategic Defense Initiative (SDI), will further decline.

By contrast with SDI, the National Institutes of Health (NIH) seem destined to flourish yet again. Last year, NIH's budget of \$6,100 million was considerably more than the administration's request. Growing concern about AIDS (acquired immune deficiency syndrome) can only loosen spending. If there is a cloud on NIH's horizon, it is the growing strength of the animal rights movement. The legislation that Congress must now take up could have a dramatic effect on the use of animals in research.

The future pattern of the US space programme should soon become a little clearer. The Challenger disaster and the two succeeding unmanned launch failures virtually crippled the programme for a time. Although launches by the National Aeronautics and Space Administration (NASA) have now resumed, many scientific payloads have been badly delayed. Space scientists will ask that NASA should augment its fleet of expendable launch vehicles, but, with an appropriation of \$2,100 million for a replacement shuttle already agreed, NASA is likely to keep most of its eggs in the shuttle basket.

Early this year, the space science board of the National Academy of Sciences will issue an important prospectus for the US

civilian space programme in the next century. The board will argue for a programme of applications contributing to human welfare. But until the shuttle flights are resumed, the present crisis can only worsen.

Elsewhere, the outlook is not so bleak. Congress has appropriated \$16.2 million to start construction of the Continuous Electron Beam Accelerator Facility at Newport News, Virginia, a welcome shot in the arm for nuclear physics.

In high-energy physics, the good news is that the Secretary of Energy, John Herrington, has declared himself in favour of seeking funds to begin work on the Superconducting Super Collider; part of the interest in the coming year will be whether he can persuade the White House Office of Management and Budget, and then Congress, to share this view. The Department of Energy will also be pushing for an international effort to build the next generation of fusion reactors, one of the issues agreed at the Reagan-Gorbachev summit at Geneva in 1985.

In the past year, the Department of

Prospects for 1987

United States

Tim Beardsley/Joseph Palca
in Washington

Energy also provided much of the impetus for one of the year's most interesting scientific debates: whether to sequence the entire human genome and, if so, when and how. Opinion seems to have hardened in favour of first making a physical and genetic map of the genome, using the interval to improve sequencing technology and the handling of the data that will emerge. The Energy Department, through its national laboratories, has already started on the mapping project, while the Howard Hughes Medical Institute is providing funds for university efforts. Serious discussions of an international effort are now under way, especially with Japan.

The Environmental Protection Agency (EPA), which has taken a strong position on the protection of the stratospheric ozone layer, will be shaping the US negotiating position at international negotiations at Geneva. Some believe that a protocol to limit emissions of chlorofluorocarbons can soon be achieved. EPA will also be concentrating on research on the greenhouse effect and development of mitigation strategies for indoor radon pollution.

One issue certain not to be resolved in 1987 is the regulation of biotechnology. The release of genetically altered organisms was a political hot potato last year, while the administration's attempts to de-



Reagan's embarrassment: admitting his decision to sell arms to Iran in his 19 November press conference.

velop a coordinated framework for regulation failed to still the many criticisms of its proposals. It will be interesting to see how many more industries and institutions choose to conduct field trials overseas.

Congress and the administration will also be paying more attention to the health of US industry, under the rubric of international competitiveness. The United States is worried that its trade deficit continues to grow, and that even its pre-eminence in high-technology products is being eroded. On the grounds of national security as well as the well-being of the economy, the government seems ready to help the semiconductor industry, which hopes to mount a collaborative manufacturing effort so as to compete more effectively with the Pacific Rim nations.

The impact on science and research of last year's Tax Reform Act remains to be assessed. Universities and charitable foundations cried foul as the new laws removed many of the incentives for charitable giving. Universities are generally concerned about the deterioration of their physical plant, while the change in federal regulations governing reimbursement for the indirect costs of research grants will force research universities to work even harder to make ends meet.

With the change of the political flavour of the Congress and the political furore at the White House, it will be some time before federal science policy is set on track. The Office of Science and Technology Policy finally has a new director, but his voice has yet to be heard in a forceful way. Erich Bloch, director of the National Science Foundation (NSF), has high goals for his agency, while it is rumoured that there is a one-third increase of funds in the federal budget for 1988, due to be released next week. If true, that could further add to NSF's importance in federally supported research. But choosing where to place research dollars will become even harder as growing demand continues to outstrip the growth of supply. □

Nakasone to tackle Japan's research funding in last year

WHAT will 1987 bring in Japan? Several of Prime Minister Yasuhiro Nakasone's major policy goals should be achieved — the break-up of the national railways, tax reform, participation of Japan in the US Strategic Defense Initiative (SDI), finalization of plans for education reform and initiation of the Human Frontiers Science Program.

The Special Council on Education Reform set up by Nakasone in 1984 will make its final recommendations this year before being wound up in the summer. But it is hard to see how the council will be able to deal with the major problems besetting Japan's higher education system — elitism in the universities and the examination 'hell' confronting university entrants — because they are simply too deeply engrained within Japanese society.

One measure that could, perhaps, have a significant influence on university education at the postgraduate level is the plan of the Ministry of Education, Science and Culture to form a graduate university to cater for the National Research Institutes for Joint Use by Universities, such as the molecular science, basic biology and physiological science institutes at Okazaki. The ministry has applied for funding for planning in fiscal year 1987.

Nobody yet knows what form the Human Frontiers Science Program will take, but it is intended to be an international research effort in basic bioscience. Sequencing of the human genome is another bioscience project that may emerge in Japan this year. The Science and Technology Agency, which is funding the development of an automatic DNA sequence analyser, has opened a 'window' of communication with the US Department of Energy and a formal agreement on information exchange is expected soon. The Frontiers programme arose in part out of criticisms that Japan takes more out of the pool of basic science than it puts in.

On 5 February, the X-ray astronomy satellite ASTRO-C, jointly developed by Japanese and British scientists, is due to be launched by the Institute of Space and Astronautical Science (ISAS) and will search for black holes and other X-ray sources. Soon after, ISAS's big brother, the National Space Development Agency (NASDA) is expected to launch the marine observation satellite MOS-1, to monitor ocean currents, water surface temperature, water vapour in the atmosphere and mineral and energy resources on land. NASDA will also launch an engineering test satellite, ETS-V, in the summer using a three-stage H-1 rocket, following the successful launch of the two-stage H-1 last year.

This year will mark increased collaboration between the United States and Japan in the marine sciences. In January, the research vessel *Natsushima* of the Japan Marine Science and Technology Center will set out on a two-month survey of the western and central Pacific as part of the joint US-Japanese study of a suspected El Niño (see *Nature* 324, 504; 1986). From April to August, US and Japanese scientists will make dives in the US submersible *Alvin* in the fore-arc and back-arc regions of the western Pacific.

Japan's best chance to make a name for itself in basic science in 1987 will probably be in high-energy physics, however. TRISTAN, the 3-km ring accelerator at the High Energy Physics Laboratory (KEK), has just undergone successful trial runs and the experimental programme should be in full swing by early May. With luck, KEK scientists may be the first to detect such elementary particles as the top quark and Higgs particle.

But it is Japan's technological feats that will continue to draw the world's admiration and ire. Construction of the new Kan-

Prospects for 1987 Japan/Australia

David Swinbanks in Tokyo
Charles Morgan in Sydney

sai international airport — a massive civil engineering project involving the construction of an artificial island off Osaka — is about to get under way at a cost of 1 million million yen (\$6,000 million).

Arguments between the United States and Japan over microchip trade seem destined to drag on interminably, and may take a strange twist — the Ministry of International Trade and Industry is contemplating accusing the US of dumping chips in Japan. But it is trade in supercomputers that seems likely to mark the next front in the high-technology war between the two world giants during the coming year.

The biggest event of 1987 will no doubt be the departure of Prime Minister Nakasone from office, which must come about before his term expires in October (unless he manages to bend party rules again). The changing of prime ministers in Japan is usually a non-event — policies carry on regardless. But Nakasone has clear-cut goals and has gone a long way towards achieving them. His implementation of tax reform and the break-up of the national railways, due to come into effect this year, may have a tremendous influence on science and technology in the long run, if they release much needed funding for the universities and research. □

Australia primes research pump

FOR the Australian government, the main priority in science throughout 1986 remained the task of linking research more closely to the needs of industry in the hope that technologically stronger industries will eventually help to solve some of Australia's economic woes.

The message of technological prosperity being preached by Minister for Science Mr Barry Jones appears to have reached the government, judging from the 1986 budget. In a tight deficit-cutting budget, science held its ground and even made real advances in some areas such as the Australian Research Grants Scheme.

According to Jones, the new era in science is symbolized by the site for the A\$18 million National Science Centre within Canberra's Parliamentary Triangle which will give science a place alongside the palaces of the arts, law and letters. Much funding for the centre has come from Japan. Australian industry has a poor record of investing in research and many of Australia's largest companies are subsidiaries of multinationals which leave innovation to other branches overseas. At the same time, the Australian home market is too small to support much research.

The 150 per cent tax deduction scheme for money invested in research and development is the government's main strategy to attract investment. The scheme finished its first full year of operation in 1986 and seems to be working. Also attracted by a cheap Australian dollar, business is expected to spend about A\$1,100 million on research in 1986-87, a doubling in real terms since 1981-82. 1986 also saw action on the recommendations of the Australian Science and Technology Council (ASTEC)'s review of the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia's largest research body. CSIRO has been encouraged to form new links with industry as a consequence of ASTEC's recommendations that CSIRO be allowed to keep any earnings arising from partnerships with industry without a corresponding reduction in its funding.

One issue that can be expected to come to a head in 1987 is the question of university funding, with the imminent release of an ASTEC report on university research. Both Prime Minister Bob Hawke and science minister Jones have spoken out on the need for more competitive funding criteria for university research. Steps have already been taken in the direction of concentrating more resources into fewer large grants to enable research programmes of particular excellence to buy major items of hardware within the Australian Research Grants Scheme. □

Year ends with Europe still divided by nationalism

EUROPEAN nationalism, which set Europe back twice this century in two world wars, could do so a third time, but this time on the economic stage. Pessimists — perhaps realists — have long signalled that the European economy is in rapid relative decline, compared not only with the United States and Japan but with all the growing economies of the Pacific, and that the currency of real economic exchange and prosperity is becoming increasingly sophisticated and technological. Meanwhile, the Pacific countries surpassed Europe by \$2,000 million as a trading partner of the United States as long ago as 1980, and the gap had already risen to \$26,000 million by 1983.

Yet European governments and administrations are still dragging their feet over the coordination of Europe's divided, inefficient, talent-wasting national research and development which should lead to new products and markets. The cancellation of the final council of research ministers' meeting of 1986 on 22 December without agreement on the European Commission's framework programme for research and development has left Europe without any coordinated plans for development. The new Belgian presidency of the decision-making European Council of Ministers thus begins this month with a huge task: to reach agreement on European research.

Perhaps Belgium, with its profound historical experience of European divisions, can forge links where others have failed. But British officials, who have struggled to reach agreement among the 12 members of the European Community for the past six months, while Britain held the presidency of the European Council, can see no easy path of compromise open to the Belgians. The issue is money, a factor of two between the 7,735-million-ECU (£5,500 million) five-year programme, a doubling of present spending, proposed by the Commission and supported, more or less, by all but Britain, France and West Germany, which would prefer no real increase in the budget. A reduced budget would certainly put a stop to the European Commission's plans to move Europe into a new phase of scientific and technological cooperation.

There will be "grave damage to Europe", says Umberto Colombo, who heads the Italian nuclear and alternative energy agency, ENEA, and is president of the Commission's scientific advisory panel and think-tank, CODEST, if the European Commission does not receive sufficient support. Colombo believes the Commission has made "a great leap forward" over the past five years in its research

programmes, and that it should now be free to move at a faster pace. In support, he singles out the successes of ESPRIT (which got information technology companies talking to one another and researching on common problems), BRITE (new technologies for old industries), RACE (telecommunications) and the Commission biotechnology and 'stimulation' (academic exchange and cooperation) programmes. "Now is the time to build on that work and achieve a genuinely European science and technology community."

But it is no secret that national interests will be paramount. Britain, for example, will oppose the Commission's plans to coordinate and share big European marine research facilities such as ships and research stations while West Germany will refuse any substantial common telecommunications programme.

The latter involves the development of a Europe-wide 'integrated digital services network' to carry mixed voice, visual and computer information. With a European market base bigger than that of the United

Prospects for 1987

Western Europe

Robert Walgate in London

States and twice that of Japan, there is a great opportunity for the European telecommunications and electronics industries with their standards and technologies to establish a solid home base and then capture world markets.

The European RACE programme in telecommunications is designed to capitalize on this position and unite administrations, technology programmes and markets. West Germany's priority however, is to promote its own electronics leader, Siemens. And with the German post office already beginning a ten-year programme to introduce integrated service exchanges throughout the country, the European Commission's RACE, a major component of the framework programme, is seen in West Germany as a means of easily tapping into German technology.

There is, however, encouragement in the form of Eureka, that *ad hoc* assembly of nominally market-oriented projects (now numbering 111 and worth £4,000 million over the next eight years) among European companies. The vitality of Eureka, which began as an initiative of President François Mitterrand of France, has surprised even Mitterrand (as he made clear in a recent speech). But it is a strange pot-pourri of projects, and there is no strong indication yet of how its typically



Next man in: Belgian prime minister Wilfried Martens, seen here voting in last October's general election, starts the year as EEC president.

bi- and tri-partite programmes will open European markets.

Without a full European framework for upstream research and downstream regulations and markets, the fear is that Eureka may flourish only in the short term — as companies allow themselves hopes of new products and wider markets, only to be disappointed in the long term if deeper issues are not solved.

One of those upstream of Eureka is the creation of a 'researchers' Europe', an efficient framework for the movement of researchers around Europe at all levels; this is part of the Commission's framework programme, an extension of the present and successful 'stimulation programme'. But ministers, while paying lip service to greater movement and cooperation, have been reluctant to fund any increase in this programme, and recently wrecked Commission plans for accelerated student exchanges (the ERASMUS programme) by voting it no money. As a result of the ministerial penny-pinching, the Commission withdrew the proposal.

Another more optimistic possibility, though, is that Eureka will help develop a completely new structure for determining European educational, research, development, technology and marketing priorities at government level. A Eureka project on high-definition television, for example, and another on the 'intelligent vehicle' (the Prometheus project for road traffic management) have both set up working groups involving governments, industry and others to develop agreements on possible national purchasing and standards. Britain has also introduced a management training programme to underpin Eureka.

Eureka may thus appear in 1987 to be beginning to take the lead in European cooperation, and, if so, government, and the Commission, will be watched very closely to see how rapidly they can respond. □

More open government — but nuclear testing on the cards £100

THE year ahead will be full of extra tasks for Soviet science, with the 70th anniversary of the Revolution likely to be an extra stimulus. The 'Complex Programme for the Integration of Research and Development' in the Comecon countries by the year 2000, announced at the end of 1985, continues. So will the tasks of implementing the Gorbachev policies of 'openness', the pruning of the academic establishment of eminent but elderly people, the reform of higher education (to make it more relevant to employment) and the elimination of bureaucratic barriers in research and development. Scientists have been told to strive not only to implement the 'scientific/technical revolution' but also to ensure a climate of peace and detente in which such progress could be achieved.

For the Soviet Union 1986 was a mixed year. The problem of the 'elimination of the consequences' of the Chernobyl disaster will spill over into the year ahead. But last year's grain harvest, announced as 210 million tonnes, was gratifyingly high — though not yet enough for self-sufficiency.

As elsewhere, acquired immune deficiency syndrome (AIDS) will be a problem during 1987. Although only isolated cases have so far been reported within Comecon, the disease can no longer be dismissed simply as a scourge of capitalism or the result of a US bacteriological warfare experiment gone tragically wrong.

Mr Gorbachev's avowed policy of openness has meant that setbacks and high-level disagreements have become public knowledge, both at home and abroad. Within the bloc, only Romania seems to stand out against the new fashion. One unexpected consequence is more frequent press reports of science-related crimes.

From the Soviet point of view, the main issue for 1987 is undoubtedly arms control, where scientists are involved in two principal ways: in opposition to President Reagan's Strategic Defense Initiative (SDI) and in support of a moratorium on nuclear tests. There has been little movement on either issue, but the Soviet media have kept up an indefatigable campaign, often involving scientists.

At the same time, while insisting that they do not intend to militarize space, Soviet scientists say they are working on 'asymmetric measures' to counter proposed US nuclear-pumped X-ray lasers. These measures (said to cost less than 5 per cent of SDI), are outlined in a recent book, *Space Weapons — A Security Dilemma*, by a team whose members all belong to the Committee of Soviet Scientists in Defence of the Nuclear Threat. Their method of 'weakening' the SDI

umbrella includes the use of missiles, space mines, shrapnel and ground-based lasers.

Meanwhile, the nuclear-test moratorium has lapsed, in the absence of an acceptable US response, as Soviet participants at the Pugwash meeting in Budapest in September had warned it might. But although the Pugwash participants said that, without a US response, Gorbachev would not be able to resist the pressure from his military hawks to resume testing, Academician Evgenii Velikhov, a leading nuclear physicist and arms control negotiator (who missed the Pugwash meeting because of illness), said in an interview on Radio Bremen in November that there is no such 'hawkish' pressure.

At home, one major decision of the past year was the cancellation of the 'project of the century', the proposed diversion south of the northward-flowing rivers of Siberia and northern Russia, supposedly on the advice of technical experts. But the writer and historian Dmitrii Ligachev, who led a campaign against the scheme, now claims its abandonment as a triumph for the "representatives of culture, the writers

Prospects for 1987

Soviet Union

Vera Rich in London

and intelligentsia". Doubts have already attached to other major hydroengineering schemes, notably the proposed Danube-Dnieper canal and the plan to transfer water from western to eastern Georgia.

In spite of this year's good grain harvest — in a year of exceptional drought — agriculture will preoccupy the planners in the year ahead. Four years after the inauguration of the grandiose 'food programme', the last major political decision of the Brezhnev era, it is tacitly accepted that the plan in its present form is unlikely to succeed. Not surprisingly, Soviet agriculturalists are seeking new solutions. Biotechnology is one answer. During Gorbachev's visit to Hungary last summer, he expressed particular interest in Hungarian achievements in this field, and selected Soviet students have been enrolled at the agricultural university of Godollo, near Budapest, which offers undergraduate courses in biotechnology.

Another solution is to encourage new forms of agricultural production. Although the new Soviet enabling legislation permitting a limited measure of private enterprise does not formally extend to agriculture (apart from small 'household plots'), some limited experiments on the Chinese model have been reported in favourable terms in the Soviet media.



These developments contrast with the situation elsewhere in the bloc. In spite of opposition from Charter-77 in Czechoslovakia and the political 'blues' and the apolitical Danube Circle in Hungary, construction of the Gabčíkovo-Nagymaros hydroelectric scheme goes ahead. In Poland, after more than four years of bureaucratic wrangling, the Roman Catholic Church has had to abandon its scheme to channel Western funds into the small private farms that account for some 75 per cent of all agricultural land.

Gorbachev's other major reforms are making slow progress. The scientific/industrial complexes so far established have failed to overcome the bureaucratic barriers between research and production, except for those such as the Paton Institute of Electric Welding of the Ukrainian Academy of Sciences which already flourished before the reform. The Soviet Academy of Sciences has a new chairman, and some leading figures in the academies of the union republics have been pensioned off, but there is still an air of gerontocracy in the academy.

The main Soviet achievement of the past year, perhaps, has been to turn the Chernobyl tragedy into a demonstration (at least after the first few days) of the new policy of openness. Soviet experts at the Vienna accident review conference more than redeemed their country's reputation. Soviet support of the Rapid Notification and Mutual Assistance conventions, and the announcement soon afterwards of an accident to a nuclear-powered submarine, substantiated Soviet assertions that the post-accident delays after Chernobyl would not occur again.

This openness, together with some Soviet relaxation on human and civil rights and the release of political prisoners in Poland, as well as partial withdrawal of troops from Afghanistan in November, have all made breaches in the barriers to the resumption of East/West scientific contacts and exchanges broken off or allowed to lapse during the political frost of the early 1980s. □

Gulbenkian Institute flourishes

SIR—Your supplement on "Science in Iberia" (*Nature* 324, 313–332; 1986) aptly describes the role of the Calouste Gulbenkian Foundation in the growth of scientific activities in Portugal. But one disturbing half-truth found its way into your text where you refer (p. 332) to the Gulbenkian Institute of Science. Although it is true that the foundation has closed down all activities in the Gulbenkian Institute of Science in areas other than biology, the institute's activities in biology are flourishing and being strengthened. The foundation has recognized that biotechnology cannot be significantly developed in Portugal without a strong research sector in basic biology. Thus, while government agencies are planning, as you correctly report, to establish an institute of biotechnology, the foundation sees an increasing need to maintain in full activity its Gulbenkian Institute of Science with respect to basic research and graduate teaching in microbiology, virology, cytogenetics and molecular genetics as well as continuing its financial support of research in other Portuguese institutions and the financing of Portuguese scientists studying abroad.

JOEL SERRÃO
HORÁCIO MENANO
J. DAVID-FERREIRA
N. VAN UDEN

*Gulbenkian Institute of Science,
Rua da Quinta Grande, Apt 14,
2781 Oeiras Codex, Portugal*

Antarctic research

SIR—The statements made by Francis Auburn in his recent review of two books about Antarctic (*Nature* 323, 22; 1986) were disconcerting to us as Antarctic scientists. We do not, of course, represent the official US government stance on the continent, but as scientists who are involved in research activities and who recently returned from a research trip there (1985–86 field season), we feel that several of Auburn's comments deserve challenge.

The National Science Foundation (NSF, which administers the US Antarctic Research Program) makes a consistent and detailed effort to educate each scientist planning to conduct a research programme in Antarctica to both the letter and the spirit of the treaty, and our interactions with researchers and support staff there indicate that most are well aware of the unique environmental and political situation in Antarctica. Most are pleased that the establishment of the treaty has allowed them to conduct exciting and unique research in this remote area. Auburn asks what NSF is trying to do in Antarctica

and mentions that its continued support of basic research should be criticized. Why? The principal thrust of the treaty was to set aside national claims and differences in order to encourage basic scientific research in and around Antarctica. He asks why, for example, the United States should support krill research when there are apparently no plans for large-scale US exploitation of this resource. But this is the nature of basic research.

In order to conduct research in the US Antarctic programme, scientists must submit proposals to NSF. Such proposals are then judged on their scientific merit, including what basic knowledge they will contribute to the field and what scientific questions they will attempt to answer. These proposals are peer-reviewed. Myopically, Auburn views this type of research as a "gift to Soviet krill exploitation", but we as scientists see basic research as a gift to all of mankind, especially future generations. From the limited perspective of today, the potential "economic" value of basic research may take decades to be realized. We believe that it is to the credit of NSF and similar organizations in other countries that they support basic research in such a remote, unique and exciting area as Antarctica. It is indeed the last frontier on this planet, and perhaps the scientific progress made possible by the treaty will serve as a catalyst in demonstrating the benefits from international cooperation.

EDITH L. SMOOT

*Department of Biology,
Hope College,
Holland, Michigan 49423, USA*

THOMAS N. TAYLOR

*Department of Botany,
Ohio State University,
Columbus, Ohio 43210, USA*

Sexist advertising

SIR—For nearly a year, *Nature* has intermittently published on the back cover an advertisement in which a woman is used in a sexually exploitive fashion (for example, the issue of 13–19 November 1986). I resent this distasteful practice and have written to you previously. Why cannot such an otherwise enlightened journal cease publishing sexist ads?

THORU PEDERSON

*Worcester Foundation
for Experimental Biology,
Shrewsbury, Massachusetts 01545 USA*

SIR—We think it high time that someone pointed out the positive aspects of *Nature's* policy on sexist advertising. You do, for example, show great restraint in not accepting advertisements such as those depicting the 'Marlboro Man'. Here, a man clad only in leather and cotton, clutching a horse, smokes a cigarette. We feel that such advertisements, which display men only as virile, macho cow-

boys, are sexist stereotypes that have little to do with reality. Men do not only ride horses. We rejoice that our male colleagues can read *Nature* every week without feeling that they have been reduced to such symbols.

Your Pharmacia advertisements (specifically, the inside cover of the 2 October 1986 issue) are to be praised. They manage to show a man and woman as hard-working, attractive, intelligent human beings. It is not even suggested that the technician knows what a horse is.

GILLIAN E. WU, SUSAN CARSON, UNA CHEN, ELLEN HSU, POLLY MATZINGER, SUSAN MCCLURE, GITTA STOCKINGER, AKIKO TAKEDA, JUDY P. WYCK
*Basel Institute for Immunology,
Grenzacherstrasse 487, Postfach,
CH-4005 Basel, Switzerland*

Mais plus ça change...

SIR—In his astonishingly accurate book *Anticipations* (of the reaction of mechanical and scientific progress upon human life and thought) (1904), H.G. Wells predicted many of the major technical discoveries with which we are now familiar.



However, he also anticipated then what would later become appropriate as a back cover for a scientific publication.

A.L. MACKAY

*Department of Crystallography,
Birkbeck College,
Malet Street, London WC1E 7HX, UK*

Simulating memory by numbers

The difficulties of making progress in neuroscience stem from the difficulty of formulating interesting (and answerable) questions. Should not associative networks be given more attention?

THERE is a sense in which neuroscience has been a great disappointment. Ten years ago, it seemed as if the question of how brains function was about to be answered by experimental investigation and analysis. There were new techniques for describing the anatomical connections of neurons and even for telling which of them were active at different times. Similarly, the catalogue of known neurotransmitter substances, already considerable, was being extended steadily while techniques such as that of 'patch-clamping' pieces of neuronal membrane seemed to offer a means of investigating the microscopic interaction between neurons and their environment in whatever degree of detail might be appropriate.

The promise of the new techniques has been sustained, yet people are not much the wiser about the working of the brain, no doubt for the familiar reason that it is easier to catalogue and describe its components than to know what questions most usefully to ask about their functioning. In the circumstances, neuroscientists should not be surprised if attempts at defining proper questions appear in unexpected places. Here is one intriguing notion, a scheme for explaining how a network of simulated neurons might be stimulated to regenerate not merely a specific pattern, likened to a memory, but a sequence of memories in predetermined order. An account of this trick, by H. Sompolinsky and I. Kanter, two physicists at the Bar-Ilan University in Israel, appeared in *Physical Review Letters* last month (57, 2861; 1986).

In reality, the new calculation is an extension of a train of rumination going back some years and most easily recognized by the label 'associative memory'. The objective is to design a network of physical entities each capable of existing in one of two distinct states (as neurons may be ON or OFF) which can simulate the behaviour of a real neural network. Those with long memories will recall that it is now thirty years since Professor Marvin Minsky launched his 'Perception', a conceptual network that might have made a working computer as well.

There are a few obvious analogies that promise that the search for a network model of the brain should not be fruitless. The hysteresis of ferromagnetic materials is a sign that strictly physical systems retain a memory of past states, while the materials known as spin-glasses, with a

virtually infinite number of distinct ground states of similar energy, are plainly capable of greater sophistication in which each equilibrium state may in principle represent a different memory.

The usual formulation of the neuronal network problem is, indeed, formally identical with that of the problem of spin-glasses. Essentially it is the creation of J.J. Hopfield of the California Institute of Technology (see *Proc. natn. Acad. Sci. U.S.A.* 79, 2554; 1982). Each node of the network is occupied by an ON/OFF entity, which in the case of a magnetic material would be physically a magnet pointing in one direction or the opposite. In principle, the neuron at each node is connected to all others in the network, which means that the properties of the network are determined by a set of coefficient T_{ij} , one for each pair of neurons i and j . If the first has no influence on the second, the corresponding coefficient is zero. If one neuron can influence another, the magnitude of the influence will be the product of the coefficient and a variable representing the state of the first neuron (mostly simply, 0 when OFF and 1 when ON).

To fix ideas, it is possible to think of the several influences on a single neuron as voltages that may be added algebraically (although more complicated schemes are permissible and, indeed, more realistic), in which case the target neuron will fire (or be ON) when the voltage exceeds some threshold. The state of each neuron is affected naturally by noise; Hopfield's model supposes that the state of each neuron at any time is nevertheless determined by a kind of instantaneous monitoring of the combination of all the influences that affect it. The neat conclusion of Hopfield's original calculation is that it is possible to choose the coefficients T_{ij} in such a way that a large number of arbitrarily chosen configurations of the system are stable against random fluctuations, or noise. The idea is that particular configurations may then be reliably evoked by suitable stimuli of the network as a whole.

Obviously the model has many attractive features. For one thing, for example, it allows that each memory state is a property of the network as a whole and not of any single element within it. For another, the model has the virtue of suggesting how a suitably configured network might be used to retrieve memories that are known, at the outset, only imperfectly; switch a subset of the neurons in the network to

represent part of a memory recalled, and the result should be to lock the rest of the network into the corresponding entire memory. In this sense, it will be noted, the brain (if truly represented by such a network) is a good deal smarter than the microcomputer now on every other desk; 'files' are retrieved not by means of an artificial filename that must be accurately remembered but by part of their own content, which may be imperfectly remembered.

What Sompolinsky and Kanter have now done is to demonstrate that if the coefficients representing the connections between pairs of neurons are asymmetrical in the sense that T_{ij} and T_{ji} are not necessarily identical, it is possible to arrange that a Hopfield network will regenerate, in sequence, a predetermined set of its stable configurations. They find it necessary also to endow the neuronal connections (called synapses, of course) with time-dependent properties. The result is that the network, when it is stimulated, will run through a sequence of its stable configurations just as if its supposed owner were recollecting a movie of some kind.

What this implies is that neurons distributed through the network are turned ON and OFF in a manner determined by the sequence of numbered patterns to which they contribute positively or otherwise. Because the representation of memory cannot be exact, but will be influenced by noise, attempts to recognize successions of remembrances by recording from single cells would probably be doomed to failure. But the authors argue well that their model would probably be even more directly applicable to the working of those neuronal networks by which animals execute rhythmic movements, running or swimming for example. In short, ways of testing such models are probably not as inaccessible as they seem.

So far as they go, these calculations are good clean fun. By themselves, they do not suggest what questions might most profitably be asked of the working of the real brains of animals or people. But there is a case for asking whether these strictly conceptual models of neuronal networks have been given the attention they deserve by those who will eventually have to test their usefulness experimentally, psychologists and others. The results so far are almost too suggestive to be untrue.

John Maddox

Optical computers

A classical finite state machine

Alan Miller

PROTOTYPE optical logic devices, which can perform the equivalent functions to transistors in electronic computers, have been available for the past few years¹ but their potential for carrying out useful signal processing and computational tasks is yet to be fully assessed (see my recent News and Views article²). Some comparisons can be made with digital electronic computers, which are now highly advanced machines comprising many logic gates (millions in large electronic systems), each of which must satisfy well-defined performance criteria. A paper on page 27 of this issue³ by Desmond Smith and co-workers from Heriot-Watt University, Edinburgh, shows that optical logic devices can fulfill the requirements of gain, cascading and restoring logic by cascading all-optical bistable devices in loop circuits while demonstrating the principle of the 'classical finite state machine'.

It is already common to superimpose information or data onto a laser beam for communication, storage or output purposes, but the data at either end of the storage or communication are still processed electronically. The challenge now being faced is how to extend the role of optics to the practical processing of data while they are still in the form of light.

Light photons have essentially complementary properties to electrons. Basic to these is the fact that electrons affect each other even at a distance, whereas light beams pass through each other unperturbed in the absence of any nonlinear interaction. This makes light a potentially useful, low cross-talk, high bandwidth means of interconnecting many devices in parallel. When one adds to these properties the fact that timing of optical pulses after propagation in free space or along fibres can be defined very precisely and that optical interactions with nonlinear materials can be extremely fast, light offers advantages precisely in the areas where the future high-speed performance of electronic computers based on small-scale integrated devices is going to be limited. Although faster electronic switching devices can be realized, their use is restricted by connectivity limitations (the number of gates or circuit units which can be interconnected with wires either on a chip or between chips), interconnection bandwidth (the data transfer rate that these wires can handle) and clock skew (the uncertainty in timing the arrival of electrical signals at different parts of a computer).

Alan Huang of AT&T Bell Laborator-

ies has discussed⁴ how light could be used to enhance computing power and, in particular, has suggested the use of optics to overcome the 'Von Neumann bottleneck' of electronic computers. In a conventional computer, for instance, there are fewer interconnections between the logic unit (central processor) and the memory than there are memory locations, so signals sent to the memory must be accompanied by addresses. Also there is a single return line for the data so the data are processed serially. The overall performance is therefore degraded when compared with the classical finite state machine in which all storage elements are accessed in parallel without the need for addresses. By using parallel beams of laser light, optics is clearly well suited to the latter type of architecture, but first it is necessary to create optical memory and logic devices that can offer cascading and restorable logic. In other words, the output of one device must be able to act as input to other similar devices while maintaining standard signals as logic levels to prevent the deterioration of information during long series of operations.

The Edinburgh group³ has developed nonlinear optical devices as three-port bistable logic elements. In other words, these are devices with two inputs and one output; for a given optical input power from a 'hold' laser beam (effectively the power supply to the device) there exist two possible transmission (or in some cases reflection) levels. The second input, a 'control' beam, can switch the device between these low and high transmission states. While the hold beam is on, it maintains the current logic level but the device is reset by interrupting this beam. These devices can function as short-term memory elements or logic gates. Signal gain between the control beam and the output is essential so that the output of one device can switch one or more subsequent devices. This demands good contrast between the low and high output levels. Restoring logic is simply achieved using control beams which are incident at an angle to the hold beam so that the output consists of only the transmitted or reflected hold beam which can be set at a standard level for all devices.

The optical circuits demonstrated use thin multilayer coatings (optical filters) deposited on flat glass plates and exhibit optical bistability when a laser beam is focused to a small spot on the plate at a specific angle of incidence. Three of these glass plates are positioned such that green beams from an argon-ion laser can be

passed from one plate to the next around in a loop using appropriate mirrors and lenses. Each plate has a hold beam incident from outside the loop to supply the required power. Because the transit time of light around the loop (about 10^{-9} seconds) is faster than the response time of present devices, the authors used a 'lock and clock' cycling system. Pulsing the input beams in the correct sequence, the bistable devices can be powered and reset so as to pass signals from device to device around the loop while carrying out a logic operation at each stage. Small numbers of parallel beams generated using holograms show that parallel arrays of data can be simultaneously and independently processed by the loop. If scaled up, these concepts could thus provide parallel processors with the relatively slow response of individual elements being compensated for by the ability to process large numbers of channels in parallel.

The new results of Smith's group³ demonstrate that repetitive manipulation and processing of parallel data streams without signal degradation is possible using all-optical devices. A processing stage (containing the computer program) would convert this into a computational machine appropriate to applications which require parallel processing of data. One example is image processing, presently carried out using computers working serially despite the fact that both the input image and the output display are in a two-dimensional form.

Much further work is required on optimization of optical elements in terms of power, speed, reliability and so on, but these devices could soon provide test-beds for new ideas on computational algorithms using architectures appropriate to large-array parallel processing. One might visualize, for instance, 1 cm^2 plates which it is reasonable to assume could dissipate 10 W of laser power. This would allow an array of 10^5 bistable elements on each plate if the holding power per element could be reduced, as projected³, to $100\text{ }\mu\text{W}$. The computing power of such devices cannot properly be calculated, however, until architectural ideas for using such a massive degree of parallelism are better developed. Some comparison may be possible here with neural networks in the brain, and their various hardware analogues. Thus, this new optical technology should not be viewed as a competitor for electronics but rather as an additional degree of freedom. □

1. Gibbs, H.M. *Optical Bistability: Controlling Light with Light* (Academic, Orlando, 1985).

2. Miller, A. *Nature* 323, 13–14 (1986).

3. Smith, S.D., Walker, A.C., Tooley, F.A.P. & Wherrett, B.S. *Nature* 325, 27–31 (1987).

4. Huang, A. *Phil. Trans. R. Soc. A* 313, 205–211 (1985).

Alan Miller is at the Royal Signals and Radar Establishment, Malvern, Worcestershire WR14 3PS, UK.

Human evolution

Out of the garden of Eden

Jim Wainscoat

A PAPER by R. L. Cann, M. Stoneking and A. C. Wilson on page 31 of this issue¹ reports that Eve was alive, well and probably living in Africa around 200,000 years ago. In considering this striking claim we should bear in mind the words of Thomas Hood in the first stanza of 'A Black Job':

The history of human-kind to trace
Since Eve — the first of dupes — our
doom unriddled,
A certain portion of the human race
Has certainly a taste for being diddled.

Cann *et al.* base their proposed identification of Eve on a detailed analysis of the mitochondrial DNA types in a diverse group of human populations. Human mitochondrial DNA, a circular molecule of about 16,500 base pairs^{2,3}, has a strictly maternal inheritance pattern. Thus, the mitochondrial DNA type of an individual is inherited from the mother, the maternal grandmother and so on.

Mitochondrial DNA evolves rapidly with a mutation rate of up to ten times higher than that of nuclear DNA⁴, giving rise to much variation between the mitochondrial DNA sequences of different individuals⁵. Cann *et al.* have determined some of the differences in mitochondrial DNA sequences from 147 individuals in 5 populations by high-resolution mapping with 12 restriction enzymes. They find 133 different DNA types and present an evolutionary tree that relates these types both to each other and to a derived ancestral mitochondrial DNA type.

The mitochondrial DNA tree has two primary roots of descent, the first leading exclusively to Africans and the second to some Africans and to all other population groups. This tree is consistent with limited nuclear DNA data that suggest⁶ an early division of populations between an African group and a Caucasian-Mongoloid group. A more difficult problem is the question of which population group is the ancestral one. The study of Cann *et al.*¹ represents the strongest molecular evidence so far in favour of the African population being ancestral. The two supporting lines of evidence are the evolutionary tree itself, which clearly suggests an African origin, and the fact that Africans seem to have more mitochondrial DNA diversity than other populations (the population which has been around longest would be expected to have more DNA sequence changes). It seems likely that modern man emerged in Africa and, as discussed in a previous News and Views article⁷, that subsequently a founder population left Africa and spread throughout Europe,

Asia and the Americas. Further studies of the mitochondrial DNA types of African populations would be valuable to provide corroboratory evidence for this 'Out of Africa' hypothesis, based by Cann *et al.* largely on studies of black Americans.

As an extension of the analysis, Cann *et al.* conclude that the mitochondrial DNAs stem from a single woman. This conclusion is not dependent on the data presented but rather is the consequence of the method of analysis. It assumes that the heterogeneity of mitochondrial DNA at the time of ancestral type was no less than in present-day populations, otherwise the present-day mitochondrial DNA types might well have descended from a large number of women who were monomorphic for a particular DNA type.

The case against the 'mother Eve' hypothesis has now been put by Latorre and colleagues, who have just published a fascinating study⁸ of the mitochondrial DNA types of the fruitfly *Drosophila subobscura*, which has recently made a dramatic expansion into the New World. Latorre *et al.* point out that in a few thousand years time all the *D. subobscura* flies may have mitochondrial DNAs derived from the type they call morph I. Clearly this does not mean that mitochondrial DNA of the descendants derives from only one *D. subobscura* currently living, as morph I is found in 44 per cent of the present-day population.

The high mutation rate of mitochondrial DNA (compared with nuclear DNA) and its inability to undergo recombinations are undoubted bonuses for the population geneticist bent on tracing ancestral lines. Nevertheless it should not be forgotten that — except for the tiny mitochondrion — the nuclear genome comprises all of our genetic makeup and stores many secrets of our murky past. For example, we inherit our mitochondrial DNA from just one of our sixteen great-great grandparents, yet this maternal ancestor has only contributed one-sixteenth of our nuclear DNA. The recently identified common female ancestor should more correctly be recognized simply as our 'mitochondrial Eve' as she has contributed little, if anything, to our nuclear DNA. Clearly, evolutionary studies of nuclear and mitochondrial DNA are complementary, although in certain circumstances they can give divergent results depending on the breeding pattern of the species.

Cann *et al.* conclude that mitochondrial Eve lived between 140,000 and 290,000 years ago. This calculation is based on an assumption of 2–4 per cent divergence in

mitochondrial DNA sequence per million years. The dating is obviously rather crude but it is tempting to relate the occurrence of the ancestral mitochondrial DNA type back to a severe constriction in population size (bottleneck). If this assumption is correct, the timing of such bottlenecks may correlate with major evolutionary events. It would be interesting to see where a more precise dating puts the ancestral mitochondrial DNA type in relation to our own evolution.

But there is an alternative explanation for the data, which is that one ancestral mitochondrial DNA type has reached fixation by random genetic drift. This is another point emphasized by Latorre *et al.*⁸ in their discussion of a population of 15,000 unrelated females which apparently has a probability of 50 per cent that the mitochondrial DNA of all individuals living 18,000 generations later will have derived from a single female⁹. This is equivalent to about 360,000 years for humans. Nevertheless, it still seems possible that the date of the mitochondrial Eve is associated with the most recent population bottleneck.

Again, these types of calculation show that many Eves have contributed to our nuclear DNA. So far, most evolutionary studies of nuclear DNA have been concerned with restriction fragment length polymorphisms and sequence comparisons from autosomal loci. However, work has now commenced on the construction of our paternal ancestry with the use of Y-chromosome-specific probes which might enable a common paternal ancestor to be identified¹⁰. It could turn out that Adam and Eve were contemporaries after all, in which case they might even have met.

Where does this leave molecular biology as a tool for studying human evolution? There are now powerful methods for studying population affinities and prehistoric migrations, but it is less clear how we can estimate the size, date and place of bottlenecks and their significance in evolution. The combined use of mitochondrial DNA, autosomal nuclear DNA and Y-chromosome-specific DNA markers has enormous potential for the genetic analysis of human populations. □

1. Cann, R. L., Stoneking, M. & Wilson, A. C. *Nature* **325**, 31–36 (1987).
2. Anderson, S. *et al.* *Nature* **290**, 457–465 (1981).
3. Barton, N. & Jones, J. S. *Nature* **306**, 317–318 (1983).
4. Vawter, L. & Brown, W. M. *Science* **234**, 194–196 (1986).
5. Johnson, M. J., Wallace, D. C., Ferris, S. D., Rattazzi, M. C. & Cavalli-Sforza, L. L. *J. molec. Evol.* **19**, 255–271 (1983).
6. Wainscoat, J. S. *et al.* *Nature* **319**, 491–493 (1986).
7. Jones, J. S. & Rouhani, S. *Nature* **319**, 449–450 (1986).
8. Latorre, A., Moya, A. & Ayala, F. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8649–8653 (1986).
9. Avise, J. C., Neigel, J. E. & Arnold, J. J. *molec. Evol.* **20**, 99–105 (1984).
10. Ngo, K. Y., Vergnaud, G., Johnson, C., Lucotte, G. & Weissenbach, J. *Am. J. Hum. Genet.* **38**, 407–418 (1986).

Jim Wainscoat is in the Department of Haematology, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK.

Molecular recognition

Feeling for the bumps

Robert Schlieff

THE problem of molecular recognition faced by the proteins that decide the fate of a bacterium infected with phage lambda is not unlike the problem of identifying a human in a dark room and then distinguishing a man from a woman. The two possible fates of the bacterium are lysis, if the phage expresses its lysis genes and multiplies until it lyses the cell; or lysogeny, if it expresses an alternative set of genes that enable it to integrate into the bacterial chromosome and replicate along with it until the lytic programme is reactivated by environmental insult. The decision between lysis and lysogeny hinges on the differential recognition by two phage proteins of operator sites in the phage DNA. Binding to these sites determines which phage proteins will not be expressed, and hence the developmental pathway followed by the phage. By a combination of reverse genetics and biochemistry, A. Hochschild and M. Ptashne and collaborators (*Cell* 44, 925-933 and 45, 681-687; 1986) have been able to show exactly how the fine discrimination required for differential binding is achieved.

The two proteins that have this pivotal role in the decision between lysis and lysogeny are Cro and repressor. Each is capable of binding to three closely related operator sites, O_R1 , O_R2 and O_R3 . In fact repressor does bind to all three operators at times in a stably lysogenized cell, and

Cro does so late in the lytic cycle. The figure shows how the binding of Cro and repressor to these sites in a newly infected cell determines the decision between lysis and lysogeny. The crucial point is that as the proteins begin to accumulate in the cell they do not bind to the three sites in the same order: Cro has the highest affinity for O_R3 and first binds there, whereas repressor first binds to O_R1 . If Cro takes the ascendancy in an infected cell, its binding first to O_R3 stops the synthesis of repressor but allows the transcription of the lysis genes. Alternatively, the binding of repressor first to O_R1 would halt the synthesis of Cro and lead to establishment of lysogeny. Thus, the choice of genetic programme followed by the phage rests on Cro and repressor proteins being able to both recognize and distinguish O_R1 and O_R3 .

The binding properties of Cro and repressor raise two questions. First, are the Cro and repressor binding sites exactly centred upon each other, or do they merely share a few nucleotides in common? Then, if the binding sites directly overlap, exactly how do Cro and repressor tell O_R1 and O_R3 apart?

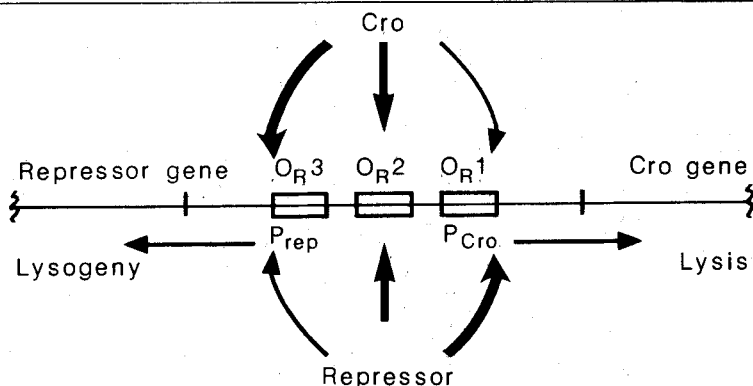
The first question is simply answered. Cro and repressor do indeed contact DNA homologously. Both proteins contact DNA primarily via a protruding alpha helix that neatly fits into the major groove of the DNA. Amino-acid residues in com-

mon between the helices of the two proteins contact common structural elements of the operators. For example, the first of the two sets of experiments from Ptashne's laboratory shows that one of these, a serine residue located at the second position in the helix of both proteins, contacts position four of O_R1 and O_R3 . This result eliminates the possibility of separate but slightly overlapping binding sites for Cro and repressor. The binding sites of the proteins exactly overlap.

The second question — how the sites are distinguished by the proteins — must therefore be answered. The most reasonable possibility, and the one which is reported in the second paper, is that the proteins directly read the three-nucleotide differences between the sites. For example, Cro detects the presence of a thymine at position three of the operator but repressor is constructed to be indifferent to the base at position three. Analogously, Cro is indifferent to the base at position eight whereas repressor contacts this position and binds significantly more tightly if it is a guanine, as is found in O_R1 . As a final demonstration, Hochschild, Ptashne and collaborators identified the amino-acid residues contacting the three nucleotides that distinguish O_R1 from O_R3 and showed that placing the Cro amino acids in repressor shifted the repressor specificity towards that of Cro and vice versa.

To demonstrate the indifference of Cro to the identity of base eight in the operators, Hochschild *et al.* synthesized operators of suitably altered sequence and measured their affinity for Cro and repressor by DNase footprinting. The demonstration that a specific residue of the proteins contacts a specific base used much the same idea. First, the affinity of the wild-type protein was shown to be sensitive to the identity of a specific base of the operator. Second, when a smaller amino-acid residue was substituted for the original residue at the position of hypothesized contact, the affinity of the altered protein became indifferent to the identity of the base in that particular position. This is as expected, for the smaller amino acid fails to contact the DNA, and therefore the binding energy of the protein is independent of the base at the position normally contacted by the substituted amino-acid residue. This type of approach has also apparently been successful in demonstrating a specific amino-acid-residue-base contact between *lac* repressor and *lac* operator (Ebright, R. *Proc. natn. Acad. Sci. U.S.A.* 83, 303-307; 1986).

What about the future? In few other systems will we have available even the relatively coarse X-ray structure data of lambda repressor bound to DNA (Anderson, J. *et al. Nature* 316, 596-601; 1985), vast amounts of biochemical data, results from nuclear magnetic resonance experiments and a rich biology, all of which



Operator sites binding repressor and Cro. Both repressor and Cro protein bind three related operators, O_R1 , O_R2 and O_R3 , but with inverse affinities (schematically indicated in the diagram by the relative thicknesses of the arrows): repressor binds O_R1 most strongly and Cro binds O_R3 most strongly. In a lysogenic bacterium, repressor is bound at O_R1 and at O_R2 ; binding of repressor at these two sites, which is cooperative, prevents transcription of Cro and other genes to the right that are necessary for the lytic cycle, but activates transcription of the repressor gene from its promoter, P_{rep} , to the left: thus repressor is continuously synthesized in concentrations high enough to enable it to bind to O_R3 , for which its affinity is relatively weak; and lysogeny is maintained. The lytic cycle is activated when an environmental insult such as ultraviolet radiation induces a protease that cleaves the repressor protein in such a way that cooperative binding to O_R1 and O_R2 is no longer possible, substantially reducing the affinity of repressor for O_R1 and O_R2 . With the release of repressor from O_R1 and O_R2 , Cro synthesis can begin from its promoter P_{Cro} . The Cro protein then binds to O_R3 and shuts off transcription of repressor so that expression of the other lysis genes is no longer prevented, and the lytic programme is initiated.

aided the experiments described here. Fortunately, this basis may not be necessary for future studies on the exact residue-base contacts between proteins and DNA. We now know how to look at high resolution for the nucleotides seen by a protein in a binding site and how to test for specific residue-nucleotide interactions. One important bottleneck to overcome, however, will be the need for predicting the precise base-residue interaction before changing the nucleotide in the binding site and the residue in the protein and then doing the biochemistry to test the prediction. This bottleneck is not severe for those proteins that use the helix-turn-helix motif for DNA recognition, because we now have an approximate idea of the geometry involved, but for other DNA-binding proteins, appropriate X-ray crystallography or more powerful biochemical

tools will be required.

To a first approximation the principle of 'feel in the right place for the right bump' describes how Cro and lambda repressor distinguish their highly similar binding sites. Nonetheless these experiments imply a deeper complexity. Both lambda repressor and Cro somewhat change their responses to the operator-specific bases with alterations in the context in which these bases are presented. Such secondary or indirect effects indicate that structure at one point affects structure at another point: that is, these protein-DNA interactions have some flexibility. It is not possible to tell whether the flexibility in this case is predominantly in the protein or in the DNA. □

Robert Schleif is in the Biochemistry Department at Brandeis University, Waltham, Massachusetts 02254, USA.

Industrial chemistry

Making money from alkanes

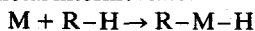
P.G. Harrison

A MAJOR objective of the petrochemical industry is the replacement of expensive olefinic and aromatic hydrocarbons by cheap alkanes, such as methane, as the primary feedstocks for catalytic conversion to high-value functionally substituted organic compounds. New catalysts and new reaction pathways are likely to be required, and there will undoubtedly be many obstacles along the way. Just how reaction pathways can be grossly modified by small changes in surface structure, for example, is elegantly illustrated by the investigations of supported rhenium catalysts by Kirlin and Gates elsewhere in this issue (*Nature* 325, 38-40; 1987).

Alkanes are available in abundant quantities from natural gas and oil deposits. The principal industrial interest lies in the conversions of the C_1 - C_4 alkanes (methane, ethane, propane and butane) to oxygen-containing compounds such as formaldehyde, methanol, acetaldehyde and ethanol, all of which are feedstocks for other processes in the chemical industry. The simple homologation of methane to higher alkanes and the dehydrogenation of C_2 - C_4 alkanes to ethene, propene, butenes and butadiene is also important. Such reactions are usually thermodynamically favourable, but there are enormous problems in transposing them into economic industrial processes. The intrinsic difficulty is to find reaction conditions severe enough to cleave the C-H (or C-C) bond of the alkane to give an 'activated' intermediate, but not drastic enough to cause exhaustive oxidation of the alkane to carbon dioxide and water. The fundamental obstacles here are the high C-H and C-C bond energies characteristic of alkanes.

This delicate problem of balance has been approached in two fundamentally different ways: activation by suitable soluble transition metal complexes in a single phase, including its recent extension, activation by naked transition metal cluster ions in the gas phase (Wise, M.B., Jacobson, D.B. & Freiser, B.S. *J. Am. Chem. Soc.* 107, 1590-1595; 1985); and activation over solid catalyst materials.

The strategy behind the single-phase (homogeneous) catalyst system involves the oxidative addition of a C-H bond of the alkane to a transition metal centre in the complex, to create a reactive hydrido (alkyl) metal intermediate:



Products are derived from the intermediate by the subsequent cleavage of the σ -bonded alkyl ligand by an appropriate reagent. Significant progress in this area has been made, particularly by Bergman, who has demonstrated the oxidative addition of simple alkanes such as methane and *n*-hexane to low oxidation state rhenium complexes to give a series of hydrido rhenium complexes such as $(C_5R_5)(PMe_3)_2Re(H)C_6H_{13}$ (Bergman, R.G. Seidler, P.F. & Wenzel, T.T. *J. Am. Chem. Soc.* 107, 4358-4359; 1985). However, the ability of such complexes to be used catalytically, that is cyclically without loss, has yet to be established. The ultimate goal of all industrial chemists — a transition metal complex which will homogeneously catalyse alkane conversions under very mild conditions — thus remains elusive.

In contrast, the heterogeneous activation of alkanes over solid catalysts has met with considerable success. Several types of

material such as metals, metals dispersed on metal oxide supports, transition metal complexes anchored on metal oxide supports, mixed metal oxides and metal phosphates have been studied, and many have proved to be catalytically active. However, heterogeneous catalysis tends to give rise to a mixture of products, and the problem is how to optimize the reaction conditions to give both high selectivity and high conversion rates. Nevertheless, several useful alkane conversions, mostly over mixed oxide catalysts, have been reported recently, including the partial oxidation of methane to methanol and/or formaldehyde by nitrous oxide over V_2O_5 - SiO_2 , MoO_3 - SiO_2 or Bi_2O_3 - SnO_2 , the oxidative dimerization of methane to ethane and ethene over lithium-promoted MgO , the partial oxidation of ethane to acetaldehyde by nitrous oxide over MoO_3 - SiO_2 , ethane dehydrogenation by oxygen over SnO_2 - P_2O_5 or chromium ions supported on silica gel, the oxidative dehydrogenation of propane to acrylic acid using oxygen over V_2O_5 - P_2O_5 - TeO_2 , and the oxidation of butane to maleic anhydride by oxygen over vanadium phosphates. Unfortunately, only fairly low conversions (often under five per cent) are obtained using methane although selectivities can be very high. Conversions of the higher alkanes are quite acceptable.

Knowing that at least some alkane conversion reactions can be accomplished over heterogeneous catalysts, the question then arises of how these systems work. The chemical reactions occur at the solid-gas interface on reaction sites which vary with the constitution of the material and its thermal history. Even the bulk solid material underlying the surface is not inert. In a heterogeneous catalyst, therefore, the reaction template is intrinsically far more complicated than that of any molecular transition metal complex, and further development will require a profound understanding of the chemistry involved in known active systems.

Many fundamental questions need to be addressed about the nature of the active material and its relationship with the surface. What are the sites for chemisorption of the alkane on the catalyst? What surface-adsorbed species are formed on the catalyst and how do they react further? What is the nature of the active oxidizing species on the catalyst? Does the catalyst suffer from reversible inhibition and poisoning?

Compared with this lack of knowledge there is a relative abundance of information concerning the use of vanadium phosphate catalysts in the oxidation of butane to maleic anhydride, where not only active phases, notably β -(VO) $_2$ P $_2$ O $_7$, but also the most important exposed crystal plane as well as sites within that plane, have been identified. The study of Kirlin and Gates, which indicates the extreme sensi-

tivity of the processes involved, represents a significant advance, not only in understanding but also in approach. These authors have shown that the nature of the catalysts derived from rhenium carbonyls anchored onto the surface of magnesium oxide can significantly affect the direction of reaction and thereby the product distribution. For example, using the supported mononuclear complex $\text{HRe}(\text{CO})_5$, only alkane hydrogenation occurs, whereas with the trinuclear cluster $\text{H}_3\text{Re}_3(\text{CO})_{12}$, both hydrogenation and hydrogenolysis,

involving C–C bond rupture, take place. Thus, data of this type allow the intelligent design of catalytically active sites able to distinguish between different possible reaction pathways, hence allowing greater product selectivity. Much more fundamental data are required in this challenging and frequently enigmatic area, but the rewards are potentially immense. □

P. G. Harrison is in the Department of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK.

Character displacement

Will grebes provide the answer?

Jared M. Diamond

CHARACTER displacement is a postulated evolutionary phenomenon in which interspecific competition causes two closely related species to be more different at sites where they live together than at sites where each lives alone¹. The differences most often discussed are anatomical, but they may also be ecological, behavioural or physiological. Character displacement might often develop during the course of speciation, and it could therefore be important — if it really occurs. But there are surprisingly few well-documented cases², most of which have been severely criticized^{3–6}. An exceptionally thorough long-term appraisal by Jon Fjelds  ^{7–12} of character displacement in the waterbirds called grebes or dabchicks, the latest instalment of which has just been published¹³, sets new standards.

Fjelds   avoided four pitfalls that have plagued previous studies. First, if one observes a single case of divergence between sympatric species, the explanation need not be competition: it could merely be chance or a point on a geographical cline arising from other causes. Hence Fjelds   considered all 20 of the world's grebe species and analysed every case in which close relatives live together or in which one species lives in isolation. Second, although most studies have examined only anatomy and only inferred ecological significance, Fjelds   directly measured diets and habitat use as well as anatomical characteristics. Third, he demonstrated the potential for natural selection by documenting how different-sized individuals of the same species compare in their dietary overlap with other species. Finally, his study was on a massive scale: he measured 2,662 specimens (from all 20 species) anatomically, examined 287 stomachs and measured 69,000 prey items.

Fjelds   identified 9 grebe populations, drawn from 4 species, that live in the absence of other grebe species. Remarkably, all these solitary populations have a 'general-purpose' bill suitable for catching

both fish and insects. All the bills are 21–26 mm long, 9–10 mm deep and 7.6–8.5 mm wide, with coefficients of variation only 4–6 per cent. In contrast, all grebe populations that are sympatric with other grebe species have bills deviating from this general-purpose model. Where two species occur together (10 areas and a total of 14 species), one population proves to be a



Hooded grebes displaying. Not until 1974 was this species discovered in remote Patagonian lakes (courtesy of J. Fjelds  ).

fine-billed feeder on small invertebrates, the other a long-billed piscivore.

In 7 cases involving 11 species, Fjelds   was able to show that bills of the same species differed between areas of sympatry and allopatry. For example, the widespread silver grebe of South America shows negligible bill variation over most of its 3,000-km range, except in Lake Junin, which it shares with a large-billed close relative (the Junin grebe) confined to that lake. Silver grebes from Lake Junin have short bills, although those from lakes only 20–30 km from Lake Junin (but lack-

ing Junin grebes) have normal-sized bills. Junin grebes, although derived from an arthropod-eating ancestor like silver grebes, have now diverged in diet and eat small fish. As another example, red-necked grebes that share the breeding range of the large-billed, fish- and squid-eating great-crested grebes have small bills and eat invertebrates, whereas red-necked grebes from outside that range have larger bills and eat fish and squid.

A plausible inference is that these populations of silver, Junin and red-necked grebes evolved altered bill size because individuals sharing a lake with another species of similar bill size found themselves eating similar food and were therefore at a disadvantage. For these three species and three others, Fjelds   tested this conclusion directly by comparing how conspecific individuals vary in bill size and diet. He found that smaller-billed individuals eat smaller prey (for example, small insects), whereas larger-billed animals eat larger prey (such as fish). In Lake Titicaca, which the long-billed Titicaca grebe and the short-billed white-tufted grebe share, shorter-billed individuals of the former have a 40 per cent overlap in diet with longer-billed individuals of the latter (to which they are similar in bill length), but the dietary overlap between longer-billed individuals of the former and shorter-billed individuals of the latter (fairly different in bill length) is only 22 per cent. Thus, individuals of two species that diverge in size do reduce their dietary overlap. In three cases (Madagascar, north New Guinea and Lake Junin), Fjelds   has tantalizing but inconclusive evidence that appreciable size divergence of dabchicks may have developed within a century of the establishment of sympatry.

Independent evidence for interspecific competition based on interference as well as exploitation comes from Fjelds  's observations of behaviour and reproductive success. On Lake Titicaca the larger Titicaca grebe attacked and chased off the smaller white-tufted grebe, but the latter moved into the former's preferred habitat in its absence. On Lake Junin the white-tufted and Junin grebes moved quickly through silver grebe habitat when the latter was present but lingered when the latter left. When the numbers of silver grebes were reduced locally by shooting, numbers of white-tufted and Junin grebes rose locally by 45 and 177 per cent, respectively, and these influxes began within 10 minutes of the removal of silver grebes. Nesting hooded grebes produced three times as many young in the absence of silver grebes as in their presence.

What are the lessons of Fjelds  's work for future tests of character displacement? Clearly, one should survey numerous species over a large area to test whether character displacement is the rule or an accident, and one should directly measure the

ecological correlates of anatomically divergent structures. Anatomical studies should focus on the structure of ecological relevance, which is surely not the same for all animals. For example, many studies have used body size, but this proves not to be the relevant variable for grebes where bill dimensions are the critical factor.

Finally, Fjelds  suggests that character displacement should be sought in stable productive habitats but not in unpredictably fluctuating ones, where food is alternately either superabundant (so that competition for food is irrelevant) or desperately short (so that any food item should be eaten). For instance, short-billed and long-billed individuals of white-tufted grebes differ considerably (90 and 60 per cent) in dietary overlap with Junin grebes in productive habitats, but there is no dif-

ference (76 and 78 per cent) in poor habitats. Thus, character displacement would be of no use to the grebes in poor habitats. The generality of this conjecture awaits testing. □

1. Brown, W.L. & Wilson, E.O. *Syst. Zool.* 5, 49 (1956).
2. Schluter, D. *et al. Science* 227, 1956 (1985).
3. Walker, T.J. *Am. Zool.* 14, 1137 (1974).
4. Grant, P.R. in *Evolutionary Biology* (eds Dobzhansky, T. *et al.*) 237 (Plenum, New York, 1975).
5. Fagerstrom, J.A. *Syst. Zool.* 27, 463 (1978).
6. Connell, J.H. *Oikos* 35, 134 (1980).
7. Fjelds , J. *Ornis Scand.* 4, 55 (1973).
8. Fjelds , J. *Vidensk. Meddr. dansk. naturh. Foren.* 136, 57 (1973).
9. Fjelds , J. *Vidensk. Meddr. dansk. naturh. Foren.* 144, 125 (1981).
10. Fjelds , J. *Steenstrupia* 7, 237 (1981).
11. Fjelds , J. *Ornis Fenn.* 59, 84 (1982).
12. Fjelds , J. *Ibis* 125, 463 (1983).
13. Fjelds , J. *Ardea* 74, 40 (1986).

Jared M. Diamond is Professor of Physiology at University of California Medical School, Los Angeles, California 90024, USA.

Thin-film technology

Diamonds at low pressure

Rustum Roy

DIAMOND remains the hardest material known, and its synthesis was for decades one of the greatest challenges to science. From Moissan's claims of having produced diamonds by freezing iron-carbon mixtures to Bridgman's systematic effort on behalf of a US industrial consortium, the goal became a benchmark. Everyone agreed that to synthesize a denser phase than the starting materials would require one to be in the thermodynamic stability field of diamond. Hence, it was reasoned, high temperature and pressure were essential to the synthesis of diamond.

The consensus, as is now well known, was wrong: diamond films several micrometres thick were synthesized at low pressure at least ten years ago in the Soviet Union and five years ago in Japan. But this development has been, inexplicably, largely ignored in the West, and industry in Europe and the United States is only now realizing its huge potential.

The unique properties of diamond show why diamond films will have applications to all sorts of abrasion-resistant coatings, from cutting tools to sunglasses and more expensive lenses. Diamond has the highest strength and hardness of any known material; a thermal conductivity 5 times

greater than copper at 300 K and 25 times greater at 80 K; extremely good corrosion resistance; maximum transparency among water-insoluble materials; and an electrical resistivity of 10^{13} Ω cm. This high thermal conductivity combined with high electrical resistivity makes diamond an ideal semiconductor substrate (see table).

When General Electric announced the successful synthesis of diamonds in 1955 an enormous range of high-pressure science began and the work provided an existence theorem for many groups throughout the world trying to develop the high-pressure route to diamond synthesis. Yet there remained those who tried repeatedly to synthesize diamond at (or below) a pressure of 1 atmosphere, where it is thermodynamically metastable only with respect to graphite. These efforts can be divided into three classes: first, the chemical vapour deposition route, where a hydrocarbon is pyrolysed, sometimes on a substrate (often one of diamond); second, a physical vapour route, via sputtering or allied processes; and third, some combination of the two, for instance, thermal pyrolysis augmented with a radio frequency or microwave plasma.

The patent by Eversole¹ was a signifi-

cant advance. Eversole claimed to have decomposed methane at a pressure of 0.1–1.0 torr at 1,000 °C and deposited a mixture of crystalline diamond and graphite onto diamond powder. Epitaxy on the pre-existing diamond structure was presented as the key that enabled the metastable diamond to form. Unfortunately, the process also produced substantial amounts of graphite, which had to be burned off in hydrogen gas at 1,000 °C for 14 hours. For two decades there were various attempts (see, for example, ref.2) to improve on Eversole's technique but growth rates remained extremely low at about 10^{-3} μ m per day.

The key discovery was by Derjaguin and co-workers at the Institute for Physical Chemistry in Moscow. Using the concept that atomic hydrogen was a 'solvent' for graphite, Derjaguin *et al.*^{3–5} described various methods for pyrolysis of methane in the presence of a high concentration of atomic hydrogen. This was achieved by various means, such as the use of an electrical discharge, a hot tungsten filament or a catalyst. The authors established by morphology and electron diffraction the formation of crystalline diamond at rates of about 1 μ m per hour.

These results did not receive the international attention they deserved. One attempt⁶ in 1964 using microwave plasma-assisted pyrolysis of methane did obtain transparent films that now appear to have been diamond and our own laboratory has recently duplicated⁷ the synthesis of large areas of thin films of crystalline diamond. Derjaguin's group is still active in this area⁸.

An important step was taken by Matsumoto *et al.*^{9,10}, who first realized the technological potential of the Soviet advances. Matsumoto *et al.* used several methods including 2.45-GHz microwave plasmas or inductively coupled radio frequency (13.56-MHz) plasmas added to a 2,000 °C tungsten filament and hydrogen/methane ratios from 500 to 100. They achieved micrometre-thick large-area epitaxial growth on (111) diamond surfaces and grew polycrystalline films of diamond on single-crystal silicon wafers, on molybdenum and on silica substrates. These results stimulated interest in diamond growth technology in many Japanese industries and universities, and dozens of papers were presented at the April 1986 meeting

Comparison of semiconductors

	Silicon	Gallium arsenide	β -Silicon carbide	Carbon
Band gap	1.1	1.43	—	5.45
Hole mobility	600	400	—	1,600
Electron mobility	1,500	8,500	—	1,900
Keyes figure of merit	1	0.456	5.8	32.2
Johnson figure of merit	1	6.9	1,137	8,206

1. Eversole, W.G. *US Patent* 3,030, 188, (1962).
2. Angus, J.C., Will, H.A. & Stanko, W.S. *J. appl. Phys.* 39, 2915–2922 (1968).
3. Derjaguin, B.V. & Fedoseev, D.V. *Growth of Diamond and Graphite From the Gas Phase* (Nauka, Moscow, 1977).
4. Fedoseev, D.V. & Derjaguin, B.V. *Dokl. Akad. Nauk. CCCP* 238, 91–94 (1978).
5. Spitsyn, B.V., Bouilov, L.L. & Derjaguin, B.V. *J. Cryst. Growth* 52, 219–226 (1981).
6. Vastola, F.J. & Wightman, J.P. *J. appl. Chem.* 14, 69–73 (1964).
7. Badzian, A. *et al. SPIE Conf. Proc.* No. 683 (1986).
8. Fedoseev, D.V., Varin, V.P. & Derjaguin, B.V. *Russian Chemical Rev.* 53(5), 435–444 (1984).
9. Kamo, M., Sato, Y., Matsumoto, S. & Setaka, N. *J. Cryst. Growth* 62, 642–644 (1983).
10. Matsumoto, S. *J. Mat. Sci. Lett.* 4, 600–602 (1985).

of the Japan Society for Applied Physics on diamond synthesis and applications.

Some of the first diamond film products from the Japanese company Sumitomo Electric include an acoustic speaker element of titanium or aluminium which is coated with about 1 μm of diamond to improve the high-frequency response. Another is the first step towards a possible diamond semiconductor, consisting of p-type 8 mm $\times$ 8 mm $\times$ 4 mm doped

yellow synthetic (high-pressure) diamonds, which can be used for substrates for the epitaxial overgrowth of an n-type layer. But perfect single-crystal p-type films and active semiconductor elements may take several years to develop. $\square$

Rustum Roy is former Evan Pugh Professor of the Solid State and Director of the Materials Research Laboratory, Pennsylvania State University, Pennsylvania 16802, USA.

Metabolism

New approaches to vitamin D

Ian Dickson

A DECADE or so ago our understanding of the main features of the vitamin D system seemed to be almost complete. But it is clear that the subject is far more complicated than was originally thought, and what seemed earlier a plausible interpretation of its role and mechanisms of action looks increasingly incompatible from the new evidence. It is proving frustratingly difficult to produce a revised hypothesis that will accommodate all these data or even to understand the reasons why symptoms of vitamin D deficiency develop in many clinical conditions. If the observations by Clements and co-workers reported on page 62 of this issue¹, which are strongly supported by a recent report from another laboratory², can be extrapolated from rats to humans, as their preliminary studies and also observations by others³ suggest, an empirical explanation for some of these disorders could be within reach.

The first significant step in understanding the mechanism of vitamin D action was the finding⁴ in 1968 that it is converted by a hydroxylation in the liver to the more active form 25-hydroxyvitamin D₃. It was soon discovered that 25-hydroxyvitamin D₃ is converted by a second hydroxylase, located in the kidney and regulated by parathyroid hormone, to 1,25-dihydroxyvitamin D₃, an even more active form which stimulates calcium absorption and synthesis of a calcium-binding protein by the intestine. This metabolite was the most potent bone-resorbing agent then known. These data were combined to produce a general explanation⁵ of the action and physiological function of vitamin D, in which the calcium-regulating hormone 1,25-dihydroxyvitamin D₃ was thought to act through the cell nucleus on two main target tissues, bone and intestine, in the manner of a steroid hormone. A fall in blood calcium would thus trigger the secretion of parathyroid hormone from the parathyroid gland which would stimulate the 1-hydroxylase in the kidney to produce more 1,25-dihydroxyvitamin D₃. The latter hormone would restore normal plasma calcium levels by stimulating calcium

absorption by the gut or by bone resorption. The classic symptom of vitamin D deficiency, the disorder of bone formation as seen in rickets or its adult form osteomalacia, was widely believed to be an indirect effect of the abnormally low plasma levels of calcium and phosphate.

We now know that the situation is more complex. Cellular binding proteins or receptors for 1,25-dihydroxyvitamin D₃ are not confined to the classical target tissues such as intestine and bone but are present in many tissues⁶ including brain, cartilage, kidney, mammary gland, muscle, ovary, placenta, parathyroid, parotid, pancreas, pituitary, skin, teeth, testes, thymus and uterus as well as macrophages, monocytes and activated lymphocytes. It may therefore be more appropriate to think of 1,25-dihydroxyvitamin D₃ as fundamental to many different types of cell, where it probably has a role in regulating intracellular calcium, than specialized to a few. It may be somewhat restrictive to think of it as a hormone that regulates calcium homeostasis as this property may be only incidental to its primary role.

Recent evidence suggests that 1,25-dihydroxyvitamin D₃ does not act only through the cell nucleus to alter gene transcription in the manner of a steroid hormone but also influences cellular metabolism by alternative, more rapid mechanisms. There are changes in the levels of certain enzymes involved in polyamine biosynthesis in duodenal cells as early as 30 minutes after administering the metabolite to rachitic chickens⁷, and calcium transport in perfused duodena from normal chicks is increased within 14 minutes of exposure to 1,25-dihydroxyvitamin D₃ (ref. 8).

The intermediary metabolism of vitamin D is also now known to be complex. More than 20 other metabolites have been identified, including 23,25-, 24,25- and 25,26-dihydroxyvitamin D₃, but their role is unknown. The biggest enigma is 24,25-dihydroxyvitamin D₃, which under normal physiological conditions is the major circulating dihydroxylated metabolite; it has

a long half-life in the circulation, suggesting that it does have some function.

Ideas about the way vitamin D acts on bone have changed dramatically in the past decade. The principal target cells for 1,25-dihydroxyvitamin D₃ in bone are now believed to be the bone-forming osteoblasts rather than the bone-resorbing osteoclasts as previously thought, and the presence of this metabolite seems to be essential for normal bone formation. The stimulatory effect of 1,25-dihydroxyvitamin D₃ on bone resorption may actually be mediated by osteoblast-derived factors⁹, but the mechanisms by which the vitamin D system influences bone formation are still unclear.

Even the term vitamin D now seems inappropriate. Normally the diet makes only a small contribution to the body's supply of vitamin D; most is synthesized from 7-dehydrocholesterol in the skin in a physiologically well-regulated process initiated by sunlight¹⁰.

Despite the remarkable advances that have already been made, there is still much to learn about vitamin D. The complexity of its actions has made it difficult to explain satisfactorily the cause of many clinical disorders. It is still not known, for example, why certain groups such as Asian immigrants to the United Kingdom and patients treated with anticonvulsant drugs should be particularly susceptible to rickets or osteomalacia.

Although the fascinating observations by Clements *et al.*¹ and by Halloran *et al.*² help to rationalize these and other clinical observations, they also seem further to expose the limitations of our understanding of the vitamin D system. The observation of Clements *et al.* that calcium deprivation increases rather than decreases hepatic inactivation of 25-hydroxyvitamin D₃, and that this is mediated by 1,25-dihydroxyvitamin D₃, is hard to reconcile with the idea that the primary role of the latter is to regulate calcium homeostasis. A more detailed investigation should help to elucidate the physiological role of the vitamin D system and the function of 1,25-dihydroxyvitamin D₃ in its many target tissues. $\square$

1. Clements, M.R., Johnson, L. & Fraser, D.R. *Nature* **325**, 62–65 (1987).
2. Halloran, B.P. *et al.* *J. clin. Invest.* **78**, 622–628 (1986).
3. Bell, N.H., Shaw, S. & Turner, R.T. *J. clin. Invest.* **74**, 1540–1544 (1984).
4. Blunt, J.W., DeLuca, H.F. & Schnoes, H.K. *Biochemistry* **7**, 3317–3322 (1968).
5. Kodicek, E. *Lancet* **i**, 325–329 (1974).
6. Norman, A.W., Roth, J. & Orci, L. *Endocr. Rev.* **3**, 331–366 (1982).
7. Shinki, T. *et al.* *J. biol. Chem.* **260**, 2185–2190 (1985).
8. Nemere, I., Yoshimoto, Y. & Norman, A.W. *Endocrinology* **115**, 1476–1483 (1984).
9. Rodan, G.A. & Martin, T.J. *Calcif. Tissue Int.* **33**, 349–351 (1981).
10. Fraser, D.R. in *Vitamins in Medicine* Vol. 1 (eds Barker, B.M. & Bender, D.A.) 4th edn 42–146 (Heinemann, London, 1980).

Ian Dickson is in the Department of Medicine at the University of Cambridge Clinical School, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK.

Quantum chaos

A beginning or an end?

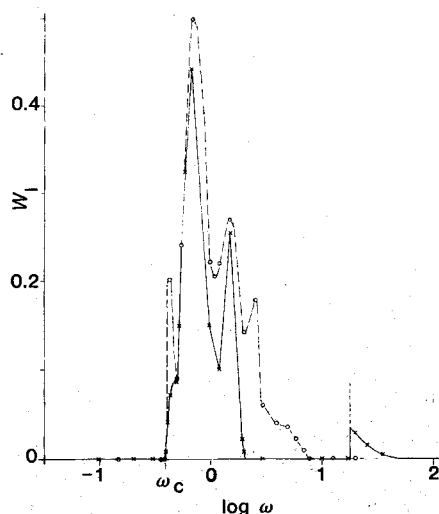
Joseph Ford

Two recent publications by Giulio Casati, Boris Chirikov, Italo Guarneri and Dima Shepelyansky (*Phys. Rev. Lett.* **53**, 2525; 1984 and **57**, 823; 1986) greatly lengthen the shadows of doubt which play across the existence of chaos in quantum mechanics. The significance of these results is but heightened by the indisputable reality of chaos in classical dynamics — games of chance, for example, as well as turbulence in all its guises. For, given that quantum mechanics is universally accepted as our most fundamental and all-inclusive description of nature, the undeniable existence of chaos in nature clearly implies that chaos must also occur in quantum mechanics. And yet the evidence exposed in these two reports indicates that it does not!

To place the new evidence in context, recall that newtonian systems are universally regarded as deterministic because their orbits have been proved to exist and to be unique. Moreover, perhaps because of a widely held, almost mystical belief that any orbit that has been proved to exist can actually be computed, newtonian systems are also presumed to be completely predictable. On the other hand, contemporary nonlinear dynamics, although affirming the notion of determinism, nonetheless establishes that deterministic chaotic motion is totally unpredictable and fully random. But what of quantum mechanics: is the quantal description of a classically chaotic system also fully random?

To understand the full import of this question, it is crucial to recall that, by definition, the term 'quantum chaos' refers only to that randomness in quantum mechanics which occurs over and above that contained in the wavefunction itself or in probabilities derived therefrom. Hence, aside from the long-debated and

still unverified possibility that chaos arises from measurement-induced collapse of the wavefunction, only two possibilities for quantum chaos exist: randomness in



A plot of classical (circles) and quantal (crosses) ionization probabilities W_i versus the logarithm of normalized microwave frequency ω for the microwave-driven hydrogenic electron obtained from numerical integration of Newton's and Schrödinger's equations. Far right is the conventional photoelectric peak. The much higher, broader and previously unsuspected peak that begins at the critical frequency (ω_c) owes its existence to chaos. Nonetheless, it is only the classical motion that is chaotic.

the deterministic Schrödinger time evolution of the wavefunction; and randomness in the deterministic eigenvalue/eigenfunction equations. It is these two possibilities that Casati *et al.* address.

Because the energy spectrum of spatially bounded, finite-particle-number, conservative quantum systems is rigorously known to be discrete, the associated wavefunctions or density matrices are almost periodic functions of time. Such predictably repetitive behaviour in the time evolution of these systems clearly precludes chaos. Consequently, Casati *et al.* have looked at finite, time-driven systems. In particular, they have numerically integrated both Newton's and Schrödinger's equations for a highly excited hydrogen atom being driven to ionization by a microwave field. Classically, they find that under suitable conditions the hydrogenic electron exhibits a truly random, diffusive absorption of energy. Quantum mechanically, the corresponding Schrödinger time evolution indicates that the electron is mimicking the classical diffusion, but more detailed inspection reveals striking deviations. First, a critical microwave

field strength has been discovered below which the quantum electron ceases its 'diffusion-like' absorption of energy even though the classical electron still continues its random walk toward the continuum. This striking numerical prediction of a quantum suppression of classical diffusion can and will be subjected to laboratory investigation. But much more damaging to the notion of quantum chaos is the fact that, even when the calculated ionization probabilities for classical and quantal electron agree quite nicely, the underlying diffusive classical motion cannot be numerically time-reversed without invoking almost 'infinite' calculational precision, whereas the apparently similar quantum motion is easily reversed using only limited accuracy. This easy time reversal clearly reveals that the quantum motion, as opposed to the classical, is not diffusive and therefore not chaotic.

In summary, despite extensive searching, no chaos has yet been found in the quantum time evolution of any classically chaotic system. In consequence, when these negative results are added to earlier ones, the shadows on the face of quantum chaos and even on quantum mechanics itself grow quite long indeed.

But it must be emphasized that Casati *et al.* also report significant positive results. Specifically, their theoretical investigations expose a previously unsuspected, quite large photoelectric ionization peak in the microwave-driven hydrogen atom. To appreciate the surprise of this result, recall that the standard photoelectric effect requires each incoming microwave photon to have an energy sufficient to raise the electron from its initial state up to the continuum. In consequence, as the microwave photon energy is decreased below that required to ionize the electron via absorption of a single photon, only unlikely multiphoton events would be expected to yield ionization.

However, these arguments do not reckon with chaos. Indeed, ionization of the hydrogenic electron was actually observed some years ago at frequencies much below those required by the conventional photoelectron effect provided that sufficiently large microwave field strengths were used. Subsequent theoretical calculations of orbits for the classical hydrogenic electron, independently performed by several investigators, then revealed this ionization to be a consequence of a transition to chaos which can be observed even at low-microwave frequency as the field strength is increased through sufficiently large values. Specifically, a critical field strength was found above which the classical hydrogenic electron absorbs energy chaotically (diffusively) and below which it absorbs energy 'almost periodically'.

Because this transition to chaos is most easily discussed using microwave frequency (or energy) rather than field strength as



100 years ago

EXTREME partisans of one school assume that the hedge-sparrow lays a blue egg because every egg that was not blue was tried in the high court of Evolution, under the clause relative to the survival of the fittest, and condemned, a hungry magpie or crow being the executioner. The extreme partisans of the other school regard the little hedge-sparrow, not only as a free agent, but as a highly intelligent one, who lays blue eggs because the inherited experience of many generations has convinced her that blue is the most suitable colour for eggs.

From *Nature* **35**, 236; 6 January 1887.

the variable, let us imagine the microwave field strength to be held fixed at some suitably large value, and then inquire what theory predicts regarding ionization as a function of frequency when each electron initially has the same energy. Note that, with this change of variable, the transition to chaos will appear at a critical frequency rather than a critical field strength.

Typical theoretical results are presented in the figure where curves of classical and quantal ionization probabilities at a fixed time are plotted against the logarithm of a normalized microwave frequency. Below the critical frequency for the transition to classical chaos, both classical and quantal calculations predict the expected negligible absorption of microwave energy by the electron. But now come the surprises numerically discovered by Casati and colleagues. First, above the critical frequency both classical and quantal curves rise steeply together and then 'level off' onto a peak much wider and higher than the conventional photoelectric peak seen on the right in the figure. Second, when these previously unsuspected ionization peaks do begin to fall,

the quantum curve drops rapidly to zero long before the classical curve does. In essence, a frequency window is predicted inside which the classical electron ionizes whereas the quantum electron does not. Because here the electron is always initiated in a highly excited, 'semi-classical' state where quantal and classical results might be expected to agree, a remarkable question now arises. Can theory predict before the experiment whether the electron perceives itself as classical or quantal in this circumstance?

In this article I have sought to convey the excitement contained in the results of Casati *et al.*, the implications of which extend from applications of the new photoelectric peak to modifications in the foundations of physics resulting from chaos. But in a deeper sense, the true excitement of these papers lies not so much in their explicit content as in the tantalizingly brief glimpse they provide of competent scientists working on problems that count. □

Joseph Ford is Regents' Professor of Physics at the Georgia Institute of Technology, Atlanta, Georgia 30332, USA.

Resonance ion spectroscopy

Counting molecules one by one

Keith Boyer

TUNABLE lasers are rapidly entering the business of identifying, counting and concentrating atoms or molecules of specific chemical or isotope species, even in the presence of large backgrounds of very similar types such as neighbouring isotopes or chemical isomers. This technique has been dubbed resonance ionization spectroscopy (RIS), and it involves the ionization of free atoms or molecules, thus converting them to a detectable form through the absorption of two or more laser quanta. One of these quanta is resonant to provide selectivity; subsequent photons, which complete the ionization process, generally have different wavelengths. Some form of mass spectrometry is often used to increase selectivity, when because of its great sensitivity, the process has been called single-atom detection. The results described at a recent meeting* show clearly that RIS is a rapidly maturing discipline of great importance.

Several outstanding accomplishments were reported at the symposium, but the most exciting was the extension of the technique to the identification and counting of both small and large complex molecules, despite the conclusion reached at the previous symposium in 1984 that molecules, as opposed to atoms, were not good candidates for the RIS technique. It

now seems that a major role of the technology will be the selective detection of molecules in science, industry and medicine.

For the application of RIS techniques to molecular species, the molecules must be prepared in a free state, and the 'hot band' rotational and vibrational contributions to their spectra eliminated. For non-volatile molecules such as chlorophyll, mesoporphrin and tripeptides, ultraviolet laser ablation can be used to introduce free molecules into a supersonic gas stream from a nozzle, thus cooling the molecules of interest and causing them to populate their lowest vibrational state (U. Bösel, Technical University, Munich). Other volatile molecules such as xylene isomers (R.J. Donovan, University of Edinburgh) and acetamide (E. Benedetti, University of Naples) can be simply evaporated into the nozzle gas stream. Acetamide is of interest as it is a basic component of peptides and proteins, whereas the ability to determine the relative amounts of different isomers of compounds such as xylene has important industrial applications.

An extraordinarily sensitive device has been developed that can detect and count 1,000 atoms in a background of 10^{15} atoms of other krypton isotopes (N. Thonnard, Atom Sciences, Oak Ridge, Tennessee). The first application of this device will be to measure the quantity of ^{85}Kr and ^{87}Kr in hydrological, oceanographical and atmos-

pheric samples to date ground water and polar ice, and to determine circulation patterns in the deep ocean or the atmosphere. Another significant achievement is the measurement, on a production basis, of impurities with a high degree of reproducibility in semiconductor substrates at the 10^{-9} level (J. Parks, Atom Sciences). Bioassays of trace elements and uranium can also be made with similar sensitivity using tiny samples of blood or urine. The medical applications are of great importance, as these are techniques which could provide rapid analysis of trace amounts of a wide variety of substances using very small samples. Small size also has the advantage of reducing the effect of reagent contamination on detection sensitivity when a chemical step is required.

The preparation of free sources of atoms for RIS measurements is being investigated both experimentally and theoretically (N. Winograd, Pennsylvania State University). Ion sputtering and electron and photodesorption of atoms indicates that both bond breaking and momentum transfer are effective mechanisms in producing free atoms from surfaces.

Various types of mass spectrometry are now being used together with RIS to improve selectivity. Descriptions were presented at the meeting of time-of-flight mass spectrometers, magnetic and electrostatic mass spectrometers and the 'Reflectron' device used to improve time-of-flight spectroscopy. Other techniques include double resonance, hyperfine structure discrimination and the use of synchronized pulsed-atom sources, detectors and lasers. Many of these techniques permit the analysis of very small samples including airborne particles, oil and blood samples, and others, allowing the detection of nanogram quantities of copper, uranium and other elements of interest. An RIS system is in use for geological and mineral exploration in China (Keling Wen, Qinghua University).

Isotope and isomer separation techniques now often involve RIS (for example, J.K. Crane, Lawrence Livermore National Laboratory). The presentations describing these techniques were limited to atomic vapour techniques to produce separations of isotopes, such as those of uranium and mercury, in large quantity although molecular techniques using flow cooling have also been demonstrated. If cost-effective, many isotopes could have great commercial value when produced in quantity in this way. One of the more exotic uses of RIS is in the search for quarks and heavy atoms, but the results to date are, unfortunately, negative. □

Keith Boyer is Emeritus Professor at the Laboratory for Atomic, Molecular and Radiation Physics, College of Liberal Arts and Sciences, University of Illinois, Box 4348, Chicago, Illinois 60680, USA.

*The Third Resonance Ionization Symposium, Swansea, UK, 7-12 September 1986. Proceedings to be published by Hilger.

Chronic granulomatous disease

SIR—Some important considerations were overlooked in the recent discussion between Creighton¹ and Orkin² concerning the putative start codon of the gene responsible for the X-linked form of the phagocytic disorder chronic granulomatous disease (X-CGD). Orkin's group has proposed that the first in-frame AUG in the X-CGD messenger RNA is not the start codon, based upon the presence of nonconsensus nucleotide (U) at the -3 position³. Creighton countered by pointing out that the first AUG in eukaryotic mRNA is the start codon in 95 per cent of known cases and by arguing that the first in-frame AUG was the more likely candidate for the start codon because of some interesting features of the inferred amino-acid sequence immediately downstream.

Creighton's points are contradictory inasmuch as the first in-frame AUG is actually the third AUG in the mRNA. (The first two AUGs are closely followed by stop codons.) In addition, further analysis of the translation initiation context strongly argues against the first in-frame AUG serving as the start codon. First, of the 211 vertebrate mRNA sequences reported by Kozak⁴ only five contain a pyrimidine (two contain Us and three Cs) at the -3 position and all have G at +4 and C at +5. Although this is a small sample, the probability that all five would contain GC at these positions by chance is ~0.0003. The first three X-CGD AUGs (including the first in-frame AUG) all contain U at the -3 position and do not have GC at +4/+5.

Through mutational analysis of the preproinsulin gene, Kozak⁵ has recently demonstrated that although U at the -3 position provides an exceptionally poor context for translation initiation, a G at the +4 position can somewhat alleviate its negative effect. Thus the lack of precedence for vertebrate start codons flanked by U(-3) and CU(+4 and +5), combined with Kozak's experimental evidence, bolsters the view that the first in-frame AUG in the X-CGD mRNA does not serve as the start codon. This example illustrates an unfortunate general feature concerning the translation start context: it is much easier to cast doubt upon putative start codons than to provide compelling evidence (both statistical and experimental) for which start codon is actually used.

DOUGLAS R. CAVENER

Department of Molecular Biology,
Vanderbilt University,
Nashville, Tennessee 37235,
USA

1. Creighton, T.E. *Nature* **324**, 21 (1986).

2. Orkin, S.H. *Nature* **324**, 21 (1986).

3. Royer-Pokora, B. et al. *Nature* **322**, 32-38 (1986).

4. Kozak, M. *Nucleic Acids Res.* **12**, 857-872 (1984).

5. Kozak, M. *Cell* **44**, 283-292 (1986).

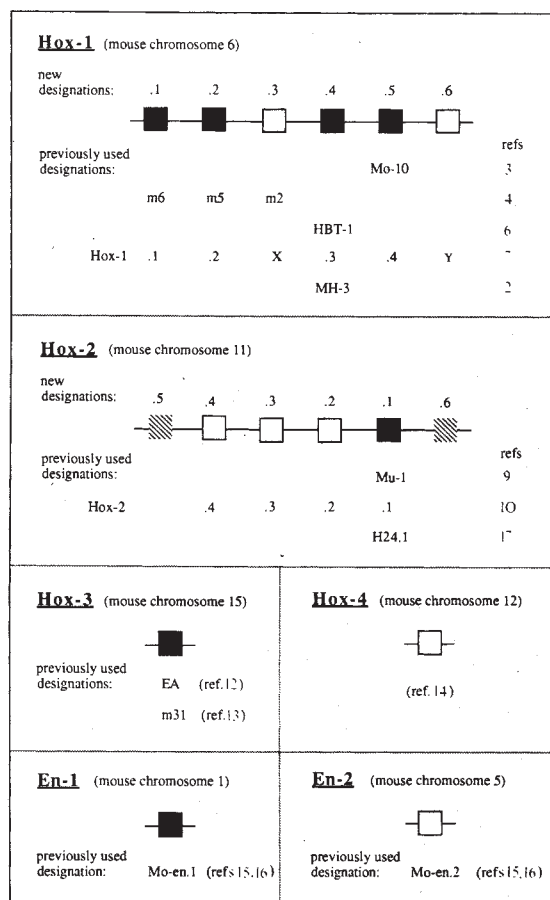
Nomenclature for homoeobox-containing genes

SIR—The recent suggestion of Gruss and Kessel¹ that a uniform system of nomenclature be adopted for the designation of the numerous mouse genes that contain homoeobox sequences was discussed at the EMBO workshop on homoeotic genes held in Leuenberg, Switzerland, 1-6 September. In accord with the rules of the International Committee for Standardized Genetic Nomenclature for Mice² the following modified suggestions are proposed: First, any homoeobox-containing gene or genomic fragment may be given the designation *Hox* (Homoeobox) providing that a significant fraction of the amino acids it encodes are identical to those of the homoeobox in the *Drosophila Antennapedia* gene. Second, the criterion for designating a new *Hox* locus is that it occupies a different map position (that is, it is physically distinct) from all other known *Hox* loci. Until this criterion is met, a new homoeobox-containing gene or genomic sequence should be designated by a 'laboratory' name. Third, the designation of any new *Hox* locus or group of loci (complex), will be determined as follows: (1), If it is not apparently closely linked to any previously described *Hox* locus, it should be numbered by accession, using the next available number in the

series (*Hox-1*, *Hox-2*, *Hox-3* and *Hox-4* have already been used to designate *Hox* loci or complexes). If two or more homoeobox-containing loci are present these will be designated by decimal numbers (for example *Hox-2.1*, *Hox-2.2*). These decimal sub-designations should, where possible, reflect the linear order of the *Hox* loci along the chromosome. (2), If the new gene is known to be closely linked to a previously designated *Hox* locus, it will be given the numerical designation of that locus or complex, and the next available decimal sub-designation in the series. Whenever possible, this should reflect the linear order of the *Hox* loci along the chromosome. (3), Before any new *Hox* designation is used in a publication, the authors are asked to contact Dr Thomas Roderick at the Jackson Laboratory, Bar Harbor, Maine 04609, USA, to ensure that this designation has not already been assigned and to inform Dr J. Peters (editor of *Mouse News Letter*) of the MRC Radiobiology Unit, Chilton, Didcot, Oxon OX11 0RD, UK, so that the symbols can be added to the annual gene list.

By applying these rules to current data, we have constructed the map shown in the figure. For completeness, we have included in this map two mouse homoeobox-containing genes, designated *En-1* and *En-2*, that have significant sequence

The filled boxes represent homoeobox sequences that have been published. The open boxes represent homoeobox sequences that have been determined, but not yet published. The cross-hatched boxes represent putative homoeoboxes that have not yet been sequenced; their existence is inferred from Southern blot hybridization analysis of cloned mouse genomic DNA fragments.



homology to the *Drosophila engrailed* gene. We recommend that this map serve as the basis for future nomenclature of the mouse genes and that this system of nomenclature be used as a prototype for naming homoeobox-containing genes or genomic sequences from other vertebrate species. When there is strong evidence to suggest that a given gene is a homologue of a specific mouse *Hox* locus, it should assume the same designation. For example, the independently isolated Hu-1 (ref. 9) and c10 (ref. 17) sequences represent the same human gene, which is the homologue of the mouse *Hox-2.1* locus, and this designation should be used in future publications describing that human gene.

GAIL R. MARTIN

Department of Anatomy,
School of Medicine,
University of California,
San Francisco,
California 94143,
USA

CO-AUTHORS: Eduardo Boncinelli, Denis Duboule, Peter Gruss, Ian Jackson, Robb Krumlauf, Peter Lonai, William McGinnis, Frank Ruddle & Debra Wolgemuth.

1. Gruss, P. & Kessel, M. *Nature* **322**, 780 (1986).
2. *Mouse News Lett.* **72**, 2 (1985).
3. McGinnis, W., Hart, C.P., Gehring, W.J. & Ruddle, F.H. *Cell* **38**, 675-680 (1984).
4. Colberg-Poley, A.M., Voss, S.D., Chowdhury, K. & Gruss, P. *Nature* **315**, 713-718 (1984).
5. Colberg-Poley, A.M. *et al.* *Cell* **43**, 39-45 (1985).
6. Wolgemuth, D.J. *et al.* *EMBO J.* **5**, 1229-1235 (1986).
7. Duboule, D., Baron, A., Mahl, P. & Gallot, B. *EMBO J.* **5**, 1973-1980 (1986).
8. Rubin, M.R. *et al.* *Science* **233**, 663-667 (1986).
9. Hauser, C.A. *et al.* *Cell* **43**, 19-28 (1985).
10. Hart, C.P. *et al.* *Cell* **43**, 9-18 (1985).
11. Jackson, I.J., Schofield, P. & Hogan, B. *Nature* **317**, 745-748 (1985).
12. Awgulewitsch, A. *et al.* *Nature* **320**, 328-335 (1986).
13. Breier, G. *et al.* *EMBO J.* **5**, 2209-2216 (1986).
14. Lonai, P. (unpublished data).
15. Joyner, A.L. *et al.* *Cell* **43**, 29-37 (1986).
16. Joyner, A.L. & Martin, G.R. *Genes Dev.* (in the press).
17. Simeone, A. *et al.* *Nature* **320**, 763-765 (1985).

SIR—As chairman of the Committee on Standardised Genetic Nomenclature for Mice I have agreed the rules proposed by Gail Martin *et al.* for naming mouse homoeobox genes and should like to endorse their suggestion that the rules are followed in other species. I should also like to suggest that the rules for naming genes in all mammalian species follow those for the mouse. The mouse and man are the two genetically best understood mammals. Which of the two is better understood is an arguable point. What is unquestionable is that genetic nomenclature for the mouse has been studied, developed and refined for far longer than that of man.

One recent point of concern is the tendency of some authors to suggest gene names beginning with a letter to denote the species in which they occur — for example *Hu-1* for a human homoeobox¹ and *Mox* for mouse homoeoboxes². If such a policy were taken seriously, all genes in man would begin with the letter H

and all in mouse with the letter M. As there are already 1,000–2,000 named loci in each species, and these numbers are expected to rise to 10⁴–10⁵, this would be remarkably inconvenient and uninformative. It is also against the nomenclature rules of both species, which state that “terminology for homologous genes should be standardised among species”^{3,4}. In all but a few cases, such as transgenic animals, the species of origin of a gene is obvious from the context.

A second point concerns the use of upper and lower case letters in gene symbols. Since genetics began, it has been traditional to use both cases of letters, with the initial letter upper case for dominant and codominant genes, and with additional use of the two cases to denote phenotypes, members of gene complexes, genes which are members of a series, and so on. Recently, human gene nomenclature has departed from this in using upper case letters throughout. This is most undesirable as it means the loss of one means of conveying information. A problem my committee faces with increasing frequency is that the amount of known information is too great to be conveyed in the symbol. Therefore, let us keep all means of conveying this information, including upper and lower case letters, numbers, hyphens and superscripts.

MARY F. LYON

MRC Radiobiology Unit,
Chilton, Didcot,
Oxon OX11 0RD, UK

1. Levine, M., Rubin, G.M. & Tjian, R. *Cell* **38**, 667-673 (1984).
2. Gruss, P. & Kessel, M. *Nature* **322**, 780 (1986).
3. *Mouse News Lett.* **72**, 3, 10 (1985).
4. Shows, T.B. *et al.* *Cytogenet. Cell Genet.* **25**, 96-116 (1979).

Micro megafaunal mounds

SIR—How big is a megafauna? Smith, Jumars and De Master (*Nature* **323**, 251; 1986) refer to “sediment mounds composed of faecal material from megafaunal deposit feeders”. I was intrigued because of my interest in Pleistocene megafaunas. I then read that the organisms referred to are echinuran worms (*Prometer benthophila*), some 55mm long. As ‘megafauna’ also refers to living or extinct large vertebrates, and ‘microfauna’ refers to vertebrates at the lower end of the vertebrate size scale, most of which are larger than echinuran worms, I wonder at the appropriateness of the terms.

The size of large cephalopods, ammonites and arthropods exceeds that of echinuran worms by a factor of ten or more, so the worms are not even in the upper size range of invertebrates. Perhaps ‘mega-invertebrate’ should describe the largest, ‘macro-’ or ‘mesoinvertebrate’ the mid-size range, and ‘microinvertebrate’ the smallest metazoans.

At least, biologists should get the vertebrate megafaunal and microfaunal cate-

gories to overlap those of the invertebrates, if the terms are to be used without any modifier; perhaps animals in the 500 mm to 2.5 m size range should constitute the mesofauna (or macrofauna).

C.S. CHURCHER

Ramsay Wright Zoological Laboratories,
University of Toronto,
Toronto, Ontario,
Canada M5S 1A1

Plant leaves may not be excreting

SIR—Ford has proposed a new function for plant leaves, that of ‘excretophore’, a tissue which serves as a means of disposal of metabolic waste products during leaf fall¹. Three common secondary plant compounds are cited as examples of metabolic waste or excretory products. The history of secondary plant metabolism from Kossel onwards² has been a progressive movement from unknown functions towards areas where a function is of suspected or, occasionally, proven significance in such fields as growth regulation (terpenoids), anti-herbivory (alkaloids and cyanogenic glycosides), phytoalexins (stilbenes) and allelopathic agents (phenolic acids)³. There is as yet no incontrovertible proof that higher plants produce any waste products other than, arguably, CO₂, O₂ and H⁺ in respiration, photosynthesis and ionic adjustment. This is precisely what might be expected of the successful autotroph.

End products of secondary metabolism must not be written off as so much waste because of ignorance of their true function. Nor is it necessary to evoke complex tissues such as leaves in which to isolate metabolic ‘unknowns’. Every mature cell in higher plants, unlike those of animal systems, is well endowed with vacuoles whose functions include the sequestration and long-term isolation of compounds which are otherwise inimicable to primary metabolic processes in the cytosol⁴. These vacuolar isolates may play a significant role in the defence of the cell against herbivores and fungal pathogens⁵. There is no sound evidence that such vacuolar compounds, which incidentally include Ford’s anthocyanins (galloyl-) tannins and oxalates, are just bursting to be excreted in one gloriously colourful autumnal flush. Otherwise one is forced to ask of a mutant which fails to complete leaf senescence⁶, is it in any way constipated?

GEORGE HENDRY

NERC Unit of Comparative Ecology,
University of Sheffield,
Sheffield S10 2TN, UK

1. Ford, B.J. *Nature* **323**, 763 (1986).
2. Quoted in Czapek, F. *Biochemie der Pflanzen* 2 Aufl. 3 (Fischer, Jena, 1921).
3. Harborne, J.B. in *Encyclopedia. pl. Physiol.* **8**, 330-402 (Springer, Berlin, 1980).
4. Matile, Ph. *New Phytol.* **105**, 1-45 (in the press).
5. Hendry, G.A.F. *New Phytol.* **102**, 239-247 (1986).
6. Hilditch, P., Thomas, H. & Rogers, J. *Planta* **167**, 146-151 (1986).

Time to leave the palace?

Duncan Davies

£92

ICI: The Company That Changed Our Lives. By Carol Kennedy.
Hutchinson: 1986. Pp.209. £12.95.

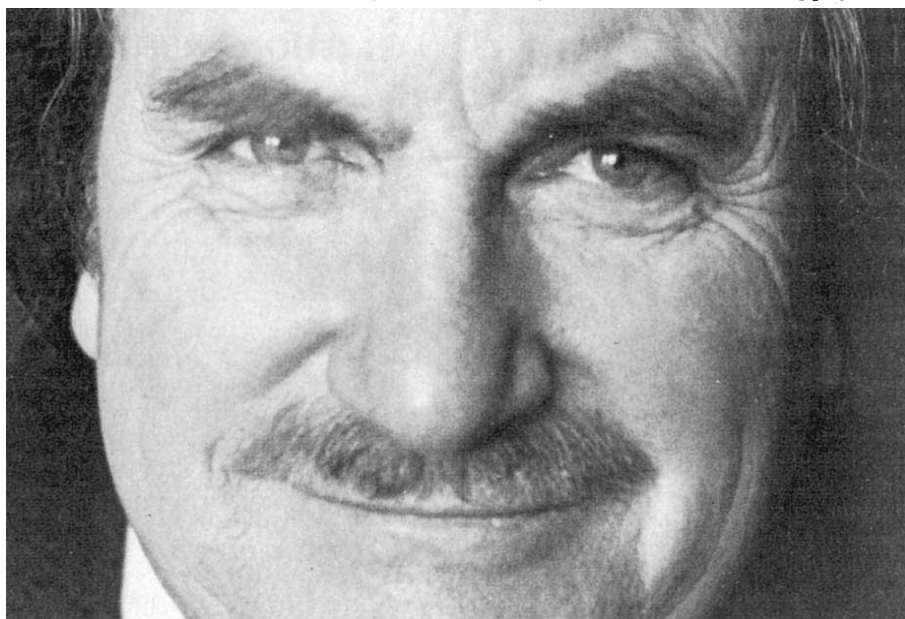
NINETEEN eighty six is an ideal year in which to write an anecdotal and journalistic storybook about ICI, to help us all understand the reasons for the success of the chemical industry during the past century. All is well with the time, the place and the loved one. If he were less modest, Sir John Harvey-Jones, the current chairman, could well say that the Firm is the true embodiment of everything that's excellent; it neither makes one sweat or squirm, for he, my Lords, embodies the Firm; The Company That Changed Our Lives, as the title has it.

First, 1986 is ICI's Diamond Jubilee. Secondly, the process of chemical invention really has transformed do-it-yourself home improvement, medicine and farming by taking systematically designed products and processes massively and profitably into craft industries. Thirdly, most of the inventors are, happily, still with us, and can tell their stories to the troubadour. Fourthly, there is a scholarly official history as an alternative source. Fifthly, the market and socio-economic conditions have been favourable to the chemical industry's talents for the whole of the past half-century, and have only just started to change. And finally, ICI itself has a record of inventive and innovative achievement that is second to none.

Carol Kennedy's treatment is racy and circumstantially persuasive, so that the book is an excellent read and should interest people who hitherto have regarded company history as prosy or incomprehensible. It was admirable to publish the book in November, and thereby include events in September, but such briskness means that there are quite a few mistakes. One of these gives a misleading picture of the reason for the British lag in the development of penicillin; others are trivial, and in total are no more than an acceptable price for a history of this kind. The harp is well-tuned and the hall has good acoustics. The photos are right.

The reader must, however, beware of the implied assumption that the commercial success in the civil market of a sequence of chemically-led inventions will continue indefinitely. Current difficulties in cashing in on innovations every bit as good as those of the past are presented as bad luck, of the kind that is inevitable from time to time. This cheerful interpretation may be wrong; it is possible that the rules are changing. Carol Kennedy's storybook has a last chapter on 2026 — the year of

the ICI centenary — which rather suggests that more of the same will continue to succeed as before. So there is no real discussion of the fact that many of the key materials for space vehicles and military systems are not finding their way into civil use anything like as readily as they used to. Large subsidies supporting new technology in agriculture have fed the hungry and are now creating very big surpluses,



Embodiment of ICI — John Henry Harvey-Jones, who retires as chairman on 31 March.

yet this is not raised as a constraint on profits dependent on further stimulation of farm yields. The decreasing rate and escalating net cost of pharmaceutical innovations are not seen to be possible indicators that new computers and biology may replace straightforward chemistry as the lead element in the next round of medical advance. The development of medical services and the pharmacopoeia may be much more closely tied together in future.

Altogether, there are strong arguments that the past chemical century is not the beginning of a thousand-year Reich. The likelihood is that from now on, in the words of an ICI Chairman, "it's going to be very different", with the chemical industry beneficially exploding into a dozen separate daughter industries that use chemistry, sometimes in a subservient role, alongside other skills. If this is so, then the fissile changes set in train by Harvey-Jones and others are going to be essential for the socially profitable exploi-

tation of the future work of Professor Birchall and his colleagues in the New Science Group of ICI.

Clues for the successful design of the agencies that in 2026 will have taken over ICI's baton lie in the fact that all organizations are the prisoners of their past successes. Study of the reasons why ICI has had to spend very large tranches of its recent pharmaceutical and agricultural earnings in supporting its polythene and polyester businesses, through processes that culminated in divestment, gives glimpses of hazards in the path of its Pharmaceutical Division. The engaging and valuable storybook that omits these points is insufficient, and is misleading if it is assumed that the sleeping princess awoken by Harvey-Jones can now live happily ever

after, in the same palace and without any special or structural reforms.

Danny Kaye once made a film called *The Court Jester*, in which there was disastrous confusion between the chalice from the palace (which contained the pellet with the poison) and the vessel with the pestle (which contained the brew that is true). We may be in the middle of a macro re-run. The palatial methods of the present chemical industry will go on earning profits for some decades to come, but well-being in 2026 will almost certainly depend on the recognition, acceptance and exploitation of quite new market and social conditions, calling for new assemblies of skills. Only then will Professor Birchall and others escape from some of their growing frustrations. □

Duncan Davies, 3 Broadlands Close, London N6 4AF, is Chairman of British Ceramics Research Ltd and President of the Society of Chemical Industry. He was the General Manager of Research and Development at ICI in the years 1969–1977.

Two minds on the solid surface £28

Michael C. Burrell & Eric Lifshin

Modern Techniques of Surface Science. By D. P. Woodruff and T. A. Delchar. Cambridge University Press: 1986. Pp. 453. £50, \$89.50.

STUDY of solid surfaces has a prominent place in a wide range of fields — heterogeneous catalysis, metallurgy, materials science and semiconductor processing, for example. Over the past 20 years such research has grown enormously, stimulated by the availability of ultra-high vacua (of about 10^{-10} torr) which have made it possible to prepare and maintain atomically clean surfaces and to investigate their structure, composition and properties.

There are now a large number of techniques for surface studies, and the aim of Woodruff and Delchar's book is to present the physical principles, strengths and limitations of each so that researchers can choose which of them, alone or in combination, will be most appropriate to their needs. In addition, recent trends in applications are discussed, giving the reader a feel for the general progress in the area concerned and its contribution to other fields.

Most of the techniques described are useful primarily in the study of well-defined, single-crystal specimens which are prepared and characterized within ultra-high vacua. For this reason, the book is addressed to those whose main interest is in fundamental surface science. Those working in applied fields will, however, also find it useful because it deals with the underlying principles on which all surface analytical techniques are based.

The book is arranged into nine chapters, the first of which briefly reviews the need for an ultra-high vacuum, the preparation of clean surfaces and the analytical methods which follow. Chapter 2 deals with surface crystallography and electron diffraction. It begins with the nomenclature used to specify surface structures, and moves on to accounts of the fundamentals of electron diffraction and (in more detail) the techniques of LEED (low energy electron diffraction). The mathematical formalisms for the extraction of surface structural information from experimental LEED patterns are introduced, and there is a succinct discussion of RHEED (reflection high energy electron diffraction).

Chapter 3 covers electron spectroscopic methods, with clear summaries of the similarities and differences between them and of the phenomena on which they are based. Chapter 4 describes techniques depending upon the interaction of ions with

surfaces, and Chapter 5 those methods based on the desorption of adsorbed surface species. Field emission and field ionization at metal surfaces are covered next, together with the associated techniques, and then (in Chapter 7) methods for measuring the surface work function and its relevance to adsorption systems. The penultimate chapter introduces the methods based on the interaction of atomic and molecular beams with surfaces, and the authors conclude with a presentation of the techniques which are used to determine the vibrational spectra of molecules adsorbed on surfaces.

The book is intended for senior undergraduates or graduates in physics or chemistry. One shortcoming for this audience is that there are no sample problems to

test the reader's understanding. Suggestions for further reading appear at the end of each chapter, however, while the figures are many and have been well chosen.

Modern Techniques of Surface Science covers all of the important techniques in current use, and is helpfully different from most of the other introductory texts on the subject in its emphasis on theory rather than applications. Further, the book is characterized by the continuity that stems from there being only two authors, and in this respect it is markedly better than similar, multi-author works. □

Michael C. Burrell and Eric Lifshin are at the Research and Development Center of the General Electric Company, PO Box 8, Schenectady, New York 12301, USA.

All about Isaac

A. Rupert Hall £26

The Newton Handbook. By Derek Gijtsen. Routledge & Kegan Paul: 1986. Pp. 665. £25. The publisher is unable to say whether or not this book is available in the United States.

TO THE making of books on Newton there is no end. But Mr Gijtsen's alphabetically arranged handbook is one of the most original in plan and not the least useful in execution. It is neither bibliography nor biography, neither exposition nor criticism, yet it contains elements of all of these and offers not only a view of Newton but of newtonianism. Mr Gijtsen will tell you what Alfred Noyes (and George Bernard Shaw) wrote about Newton; whether Newton was chaste or not, and his niece too; how many portraits of Newton were painted, and what fluxions or Leucateello's balsam are. I am only sorry that he has no separate alphabetical entry — there are some 550 in all — for *Ballistics*, or *Diana's Doves* or *Malebranche*.

Aficionados may easily develop the devising of yet more headings into a parlour game. *Diamond* is there ("One of the most often repeated anecdotes about Newton...") but not the story of Newton and his cats. *Apples and Apple Trees* do earn a heading: sadly, the most famous apple since Eve's, a Flower of Kent, is described as flavourless (and it is not bright red). Most of the things one needs to know about Newton's life and writings are to be found in the book, neatly compressed, and I for one mean to have this useful work of reference always to hand.

Mr Gijtsen has thoroughly worked through the extensive newtonian editions, books and articles of recent years and includes notes on the more notable authors of them. He has also made good use of the invaluable bibliography by Peter and

Ruth Wallis. He is therefore up-to-date in analysis and interpretation, and is well informed about the manuscript sources from which so much has been harvested.

Inevitably, because of the vast amount of material to be sifted, summarized and arranged there are errors. The presentation of mathematics is not invariably happy (examples pp.150–151 and 214–215); it is misleading (p.215) to omit the late date (early 1690s) of Newton's dot notation for fluxions, correctly noted on p.338. Less seriously, the ownership of a ton and a half of gold and a quantity of silver appears to be attributed to Newton at absurd valuations (p.277): what is at stake is the Mint Master's percentage on these weights of precious metal. Sometimes proper names, such as those of Michael Sendivogius and Valentine Greatrakes are improperly rendered; and to write of "Classical physicists..." (p.269), when Greek philosophers are meant, provokes a shudder. The treatment in various places of newtonian matter-theory is curiously inadequate and Stephen Hales is wholly omitted. The reprints of Hessen's notorious essay on the economic interpretation of the *Principia* (1931) might have been recorded.

One could go on, but though mistakes of this sort suggest that cross-checking to the original sources is prudent, they do not seriously weaken one's sense that Mr Gijtsen is a well-informed and trustworthy cicerone to the complexities of newtonian scholarship. I have found his summaries of the present state of scholarly thinking accurate and on the whole consistent. With a little patience (as regards topics, the index could be richer) the reader can find his way into any theme, and discover the primary and secondary sources bearing upon it. More one cannot ask. □

A. Rupert Hall, 14 Ball Lane, Tackley, Oxford OX5 3AG, UK, is Emeritus Professor of the History of Science and Technology at the University of London.

Players in policy

Alastair Hay £41

Pesticides in World Agriculture: The Politics of International Regulation. By Robert Boardman. Macmillan, London: 1986. Pp. 221. £27.50.

Books about international regulation tend to fall into one of two categories. They either describe the management of events behind the scenes by the big time players or they serve as maps, guiding the untutored, or vaguely informed, along the by-ways of international bureaucracy.

Pesticides in World Agriculture is in part a road-map, describing the United Nations agencies — such as the Food and Agriculture Organization, World Health Organization, United Nations Environmental Programme and International Labour Organization — which are involved in the formulation of international policy about pesticide manufacture and usage. But in his accounts of the activities of the agencies Boardman goes further than simply giving place names and mileages, and he succeeds in explaining how the UN bodies became involved with chemical pesticides in the first place, and why particular agencies act as they do.

Pesticides first became an item on international agendas in the 1950s, and it was at that time that FAO and WHO started to take an interest in them. The differing roles of the two agencies help to explain some of their actions. The business of FAO is to maximize food production. It therefore has a closer working relationship with the agrochemicals industry than does WHO which, through its remit on health, runs the international programme on chemical safety.

The two agencies are co-sponsors of the renowned *Codex Alimentarius* Commission, which documents and pronounces upon pesticide residues in food. *Codex* publications are required reading for anyone concerned with residues; they set out certain permissible levels and explain how these have been determined. To reach a decision the Commission is dependent on industry for its data, and it comes as no surprise to find that many companies are not particularly forthcoming in this respect. Having spent vast sums of money developing and testing a chemical, they are reluctant to trust an international agency with this privileged material. Boardman shows that national regulatory bodies, particularly those in the developing world, also suffer from this problem. Further, while North American, European and Japanese government agencies are usually able to get the information they need (if necessary by resorting to the courts), they are not always as cooperative as they might be in passing on the informa-

tion to their counterparts in Brussels, Rome and Geneva. Companies, it seems, are not the only guilty parties.

Reconciling the competing demands of environmentalists and chemical companies is an impossible task, and the best that can usually be achieved is a sort of uneasy truce. This is where the international agencies come into their own; they are seen by many to be above the fray and thus in a position to perform the role of international arbiter. But this function is not always carried out as impartially as we are led to believe. Lobbyists for particular positions are less visible at the international level, but they are no less active. Even WHO, which keeps its distance from industrial interests, is not immune, as the downgrading of its campaign against the evils of tobacco attests.

Nevertheless, it is vital to have some forum, however imperfect, where the benefits and drawbacks of pesticides can be considered in a global context. Those offered by FAO, and by the Organization for Economic Cooperation and Development (OECD), enable this to be done. Boardman points out that it is in these organizations that the different needs for pesticides in temperate and tropical climates can be discussed, as well as the effects of humidity, temperature and soil conditions on the rate of breakdown of the chemicals. Moreover, decisions reached by international bodies often strengthen the hand of the national agencies. Boardman notes that OECD guidelines — to which both government and co-opted industry specialists are party — often aid the US Environmental Protection Agency in formulating domestic legislation.

To illustrate the diversity of regulatory problems, the author describes some case-histories, among them that of the herbicide 2,4,5-T and its contaminant dioxin, and the attempts at regional harmonization of policy such as that performed by the European Economic Community. 2,4,5-T has become an international symbol for environmentalists and the chemical industry alike. The former claim that it illustrates the perfidy of the agrochemical trade, the latter that the furore shows the political rather than scientific nature of debates about pesticides, even those with-

in the purview of the EEC. For administrators trying to anticipate trans-national decisions about pesticides, and what the feedback effect on domestic regulations might be, it must be like playing three-dimensional chess against a grandmaster.

Boardman's last case-history deals with the use of pesticides in Malaysia. If ever there was a need for regulation, this section of his book illustrates it; one survey found that 71 per cent of farmers in the country were unaware of the toxicity of pesticides. Small wonder then that 120 people should be poisoned by endosulfan in an incident in Sarawak, or that five members of one family in Iban should die from paraquat residues on rice.

Over 100 companies are registered in Malaysia as pesticide manufacturers or formulators, and clearly the industry has a strong foothold. For all that, Malaysia is probably one of the better nations in the developing world in keeping tabs on pesticides. The environmental groups are powerful and have, for example, forced the government to look into the adequacy of warning labels on pesticide containers. Also, the two sets of interests are not invariably in conflict; the adulteration of pesticides is a common problem, and the big chemical companies have at times found it helpful to support regulation, for this in turn will protect their own products.

Malaysia has used the international agencies, FAO in particular, to help it formulate its policy. But, as Boardman points out, without enough good administrators and enforcement officers the government has had to build on a "slowly-emerging consensus" about the control of pesticides, and has been subject to criticism for its gradualist approach. He believes that Malaysia's policy is appropriate and is successful; how successful, only subsequent statistics on poisoning episodes will show. What is clear is that in the field of pesticide regulation individual scientists, countries and regulatory agencies are all bit players in a larger drama. This book is not easy reading, but it will help its readers get to the theatre. □

Alastair Hay is a Lecturer in the Department of Chemical Pathology, University of Leeds, Old Medical School, Leeds LS2 9JT, UK.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Chemical control — a crop of oil seed rape is sprayed with herbicide and fungicide.

Embryonic events

Jonathan Cooke £28

Developmental Biology: A Comprehensive Synthesis. Vol. 2 The Cellular Basis of Morphogenesis. Edited by Leon W. Browder. Plenum: 1986. Pp. 651. \$79.50, £66.95.

LIKE many other areas of science, developmental biology supports a dual publishing industry — the slim volume of ideas and the hefty, multi-author tome of information. The genre in which single authors marshal information to underpin particular ideas is currently, perhaps permanently, out of favour. That may justly reflect the true nature of the beast, i.e. the embryo. Certainly, in referring on average to 140 papers and reviews apiece, the present contributors make their 16 chapters tough going for anyone in search of what the title suggests, a synthetic view.

The brief for the contributors to this, the second volume in the series *Developmental Biology*, was to ask whether the architectural and shape changes in embryos can be understood as the controlled deployment of a restricted repertoire of cellular mechano-adhesive behaviours — or of their molecular bases, if you prefer that level. Some of the authors seem to have long since lost sight of the original enterprise, finding satisfaction merely in the collation of information interlarded with speculation over parochial matters. There are chapters on endogenous lectins, on sea urchins and on early mammalian development that would surely have delighted Diderot, though I cannot believe that encyclopaedism was their authors' intention. Moreover, the central, meatiest chapters are grouped under the hoary rubric of "Cell-Cell Interactions in Development" (CCID!), which in its generality transcends the book's title and attests that, increasingly, editors of such projects pin down contributors in what order they may and invent intellectual frameworks as they go on. But if you belong to the growing number of biologists who assimilate and thrive on information *per se*, you could treat yourself here. And as a bonus, a few chapters rise above the blandness with a level of individual excellence.

The unsatisfactory aspects of the book are foreshadowed in the opening chapter on cytoskeletal dynamics and are confirmed in those on subsystems modulating cell adhesion. These, after all, are the very nuts and bolts for all that is discussed elsewhere, but the authors retreat — sometimes vaguely downcast but elsewhere apparently happy — from proposing any unifying syntax of events among these structures that links them systematically to the behaviours we see in cell groups. One trouble is that the subsistence of cell

adhesiology has been in a mess since the acronym CAM began its proliferation. Some of the cell-adhesion-influencing molecules characterized *in vitro* do look exciting. But we have little idea whether cells summate the effects of individual components such as these, or ignore all but a crucial one, when performing discriminatory behaviours *in vivo*. Yet, as emerges clearly from the discussion of the diverse cell populations that de-epithelialize and migrate, there is a commonality of molecular repertoire — so much so that avoidable redundancy is introduced into the book in several, perhaps under-edited, chapters that appear in succession.

The inseparability of pattern — the often finely tuned positional specificity of cells' mechanical behaviour — from mor-

phogenesis rears its troublesome head in at least two contributions. The authors concerned admit that we have little idea even as to which parts of the cell's machinery underlie the positional modulation of such behaviour, as indeed of later differentiation choices.

I was disturbed to find Keller's chapter on the mechanics of gastrulation, over-long and dense as it is, so relatively engrossing and admirable. Partiality due to my own intimate research-level knowledge? I think not. At his chosen explanatory level, but with due reference to research at levels "above" and "below", he is here pre-eminent in the pursuit of overall understanding in one whole system. □

Jonathan Cooke is at the National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.

Observing life

A.J. Meadows £26

Of Stars and Men: Reminiscences of an Astronomer. By Zdeněk Kopal. Adam Hilger: 1986. Pp. 489. £29.50, \$60.

BEYOND doubt, they do not make astronomers the way that they used to. Thirty or forty years ago, British astronomers typically worked on their own. In Britain, theory still dominated observation, as it had done for decades before, and entrants to the profession were still expected to be competent mathematicians. Compare this with the present position, where research groups are common, most astronomers have a background in physics and the British effort devoted to observation exceeds that devoted to theory. The main cause for this switch was the advent of the space age at the end of the 1950s, which brought a flood of newcomers into astronomy with new backgrounds and interests. But the space age was itself one aspect of a wider change that has affected astronomy since the Second World War — the development of new tools for studying the Universe. The first of these was radio-astronomy, in which Manchester and Cambridge led the way in Britain.

Professor Kopal's period of office in the Chair of Astronomy at Manchester stretched for 30 years from the early 1950s, so covering the transition period in British astronomy. In his early days there, most of his postgraduate lecturing was done at Jodrell Bank, but, as time passed, the Astronomy Department itself gradually expanded. Anyone involved in British astronomy during this period will recognize familiar developments in Kopal's story. For example, the lunar mapping project at Manchester, with which Kopal's name was associated during the early days of the space age, is almost a

prototype of the group research activity now firmly established in astronomy.

Kopal's original astronomical training in pre-war Czechoslovakia was along traditional lines, but he was also an enthusiastic observer from an early age. With typical initiative, he financed his first major observing trip — at the age of 22 to view a solar eclipse from Japan — by translating and publishing a Czech version of James Jeans's *The Mysterious Universe*. He became a frequent and enthusiastic traveller; yet most of his working life was spent in two places, first Harvard and then Manchester. At both, he was known for his willingness to speak his mind. Indeed, when he came to Britain he was particularly known as the man who had clashed with the then doyen of American astronomy, H.N. Russell. Kopal devotes several pages to this argument along with his (somewhat ambivalent) impressions of Russell. More generally, he was obviously unhappy with post-war developments in East Coast astronomy. He comments at length on what he sees as the downhill path of the Harvard College Observatory after Harlow Shapley retired as director. Nor were his days in Britain totally without controversy. For example, he rehearses at some length the debate that surrounded the setting up of a new journal, *Astrophysics and Space Science*, with himself as editor in the late 1960s.

Though Kopal is a firm proponent of modern astrophysics, he retains something of the scholarly tradition of an earlier era, when professors were expected to be able to talk about almost anything. In consequence, his story wanders off into many interesting by-ways: but the main theme — the rise of modern astronomy — shines clearly throughout. □

A.J. Meadows, formerly Head of the Departments of Astronomy and History of Science at the University of Leicester, is now a Professor in the Department of Library and Information Studies, Loughborough University of Technology, Loughborough LE11 3TU, UK.

The demonstration of restoring digital optical logic

S. D. Smith, A. C. Walker, F. A. P. Tooley & B. S. Wherrett

Department of Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK

Nonlinear all-optical circuit elements can be used to create indefinitely extensible optical logic. By the use of a 'lock and clock' architecture and an off-axis configuration of the power and signal beams, signal corruption can be avoided and standardized signal levels through cascaded devices maintained. Several loop circuits have been operated to demonstrate a simple 'optical classical finite state machine'.

ALL-OPTICAL nonlinear signal processing devices based on the principles of 'optical bistability' can be configured, by setting initial conditions, to give a family of optical circuit elements¹⁻³. These devices, which operate by a combination of nonlinearity and optical feedback, exhibit a hysteretic transmission/reflection response to changing optical input within a controllable range of incident powers (see Fig. 1). This bistable region is bounded by two discontinuities at which switching between the two states can occur. Such a bistable optical switch can act as a write/read/erase binary memory element^{4,5}. Alternatively, by reducing the bistable range to zero or nearly zero, non-latching logic gates can be realized. The use of transmission or reflection outputs, combined with appropriate steady-state optical bias ('holding power') makes the following binary logic functions available: AND, OR, NAND, and NOR⁶.

An essential condition for any system of restoring digital logic is that each logic gate must be able to drive at least one other similar gate—and this implies a region of gain between the valid logic-zero and valid logic-one states⁷. Transistors fulfil this requirement very successfully. But all-electronic devices require a 'settling time', as capacitances are charged, before the next logic cycle can proceed. Electrical interconnections have resistance-capacitance (RC) time-constants that cannot be reduced by scaling to smaller size⁸. This communications 'brick wall' will, in the future, set a limit to the speed of operation of very-large-scale-integrated (VLSI) circuitry.

The communications possibilities of optical signals—literally speed-of-light transfer at high bandwidth—will avoid this RC time-constant limitation. Furthermore, optics can be exploited to give unrestricted interconnections between large numbers of points, in parallel⁹ ($\geq 10^7$), and with extremely low time skew (≤ 1 ps)—a performance inconceivable with electronic interconnections. In this article we show how the natural storage capability of bistable optical switches can be exploited to propagate binary signals around a digital optical circuit.

The simple loop circuit, described herein, constitutes a first 'optical classical finite state machine'¹⁰ in which the parallelism of optical devices allows simultaneous memory interrogation (Fig. 2a). This is in contrast to the conventional Von Neumann machine where memory is accessed, via an address mechanism, one word at a time (Fig. 2b).

Optical circuit elements

During the European Joint Optical Bistability (EJOB) project we have fabricated and tested a variety of optical circuit elements by using both refractive and absorptive nonlinearities in a range of materials and both electronic-excitation and thermal mechanisms have been exploited⁶. To the computer scientist, only the functional performance is significant—power, speed, stability and so on. Whereas devices operating with low optical powers and based on fast nonlinearities associated with electronic transitions would be ideal, results so far show that laser-induced heating coupled with a large change of refractive index with temperature, dn/dT , gives a first generation of usable devices

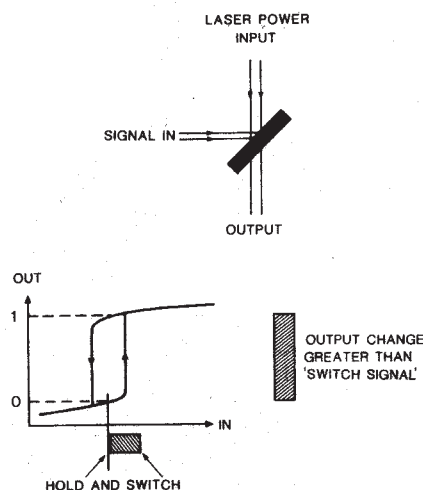


Fig. 1 An all-optical OR gate with an off-axis signal input capable of switching the transmission state encountered by the higher-power holding beam. By holding close to the switch point the relatively small signal can be digitally amplified as the device switches between states.

that fulfil most of the requirements of initial digital optic experiments. At first sight it might be thought that thermal devices will be slow, but use of thin film technology can give device volumes of a few cubic micrometres and we have shown both theoretically¹¹ and experimentally¹² that response times can be in the nanosecond region, although more typically turn-on and turn-off times are in the $1\ \mu\text{s}$ – $1\ \text{ms}$ range. We concentrate on such thin-film devices: nonlinear interference filters (NLIF).

The simplest form of narrow-band interference filter consists of a spacer layer, of integral number of half waves optical thickness, of a high index material surrounded by quarter-wave stacks (alternating low/high refractive index), all deposited on a substrate (for example, glass, G). We may thus describe the general design as $G[m(\text{HL})][n(\text{HH})][m(\text{LH})]$, in terms of quarter-wave optical thicknesses of high (H) or low (L) index material. A typical structure may contain about 17 layers. Our currently preferred materials are ThF_4 for the low-index layers ($n = 1.4$) and either ZnSe or ZnS for the high index ($n \approx 2.5$). These are deposited by thermal evaporation *in vacuo*. The total physical thickness of the multilayer structures is typically $\sim 1\ \mu\text{m}$.

The process responsible for optically bistable switching in these devices is as follows. The interference filter has a narrow band of transmission ($\sim 4\ \text{nm}$ full width at half maximum (FWHM)) centred just on the short-wavelength side of the 514-nm emission of an argon-ion laser. The absorption properties of the spacer layer are such that sufficient laser power is absorbed from a $\sim 100\text{-mW}$ optical input, over a $100\text{-}\mu\text{m}$

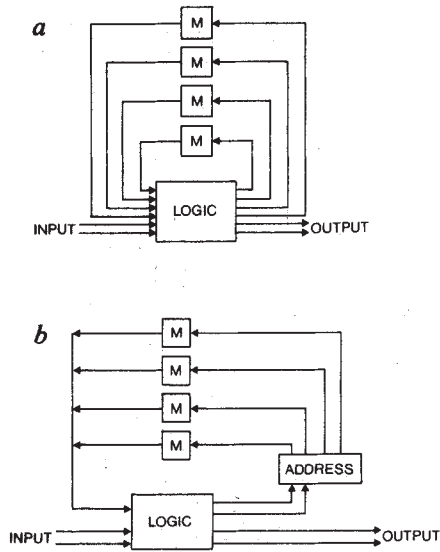


Fig. 2 *a*, A classical finite-state machine, with parallel access to all memory elements. *b*, A modified finite-state (Von Neumann) machine, with only serial access to memory via an address. The reduced number of connections is achieved at the expense of creating a time bottleneck.

diameter, to cause an increase of temperature of ~ 50 K. This shifts the absorption edge, and hence changes the refractive index enough (by $\sim 1\%$) to tune the filter into resonance with the laser line. It is the *internal irradiance* (I_{int}) that is effective in this process and this is greatest in the *spacer* layer—increasing strongly as resonance is approached according to (assuming a low internal absorption)

$$I_{\text{int}} = I_i T(\lambda) (1 + R) / (1 - R) \quad (1)$$

where I_i is the incident irradiance, R is the effective reflectivity of the stacks and $T(\lambda)$ the transmission of the filter, sharply dependent on the proximity to resonance of the device. Extending the Airy equation, which determines $T(\lambda)$, to include the nonlinear phase change, $\phi(I_{\text{int}})$, we have

$$T = \frac{1}{1 + F \sin^2(2\pi nd/\lambda + \phi(I_{\text{int}}))} \quad (2)$$

where nd is the optical thickness of the spacer and F is the coefficient of finesse.

Equations (1) and (2) describe the switching behaviour of the device assuming a simple plane-wave approximation. More sophisticated multilayer matrix-coupled parameter theory has been developed¹³ for detailed design of such optically bistable filters—based on absorption-driven refractive nonlinearity.

The actual temperature rise of the illuminated portion of the filter (ΔT) depends upon the spot dimensions and the heat flow into the substrate. If the multilayer structure is thin and the illuminating spot radius (r_0) is relatively large, then the thermal response is dominated by the properties of the heat-sinking substrate: thermal conductivity, κ_s ; specific heat, C_p and density, ρ . Thus ΔT and the limiting switch time constant, τ_0 , are given by¹⁴:

$$\Delta T = P_0 A / 2\sqrt{\pi \kappa_s r_0} \quad (3)$$

and,

$$\tau_0 = r_0^2 C_p \rho / 4 \kappa_s \quad (4)$$

where P_0 is the incident power and A the fractional absorptance of the film. It can be seen from (3) and (4), and is confirmed experimentally^{15,16}, that at large spot diameters, switch power

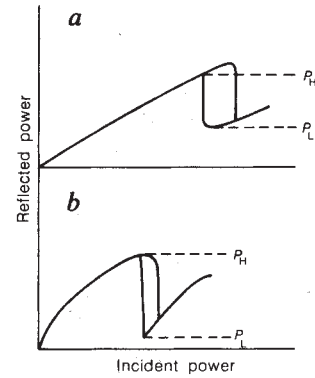


Fig. 3 Reflection-output transfer characteristics of ZnSe NLIFs. *a*, Contrast (P_H/P_L) = 2.24; *b*, P_H/P_L = 9.5. The enhancement of the contrast (*b*) was achieved by increasing the reflectivity of the back multilayer stack and observing only the uniformly illuminated region. Arbitrary units used.

is proportional to radius and the switching time is proportional to area.

For a spot size of $10 \mu\text{m}$ diameter, switch powers can be ≤ 10 mW. The total switch speed is strongly influenced by critical-slowness effects¹⁵ but has a well-defined lower limit determined by the thermal time constant of the device, τ_0 . For a $10 \mu\text{m}$ spot this is $\sim 10 \mu\text{s}$. This gives a characteristic switch energy of 100 nJ. By physically pixellating the device to inhibit transverse thermal diffusion, increasing the spacer thickness and optimizing the distribution of absorbed power we predict that this switch energy can be reduced to 1 nJ (ref. 17), assuming the same $100 \mu\text{m}^2$ active area. The use of device areas of a few square micrometres should further reduce the switch energy to ~ 10 pJ. Interestingly, this is comparable to the switch energies of other (room temperature, visible/near-infrared) all-optical bistable devices, including those based on electron-transition nonlinearities. Multiple-quantum-well (MQW) GaAs/GaAlAs, for example, requires $\sim 10^5 \text{ W cm}^{-2}$ with 20 – 40 ns switch times¹⁸ or $3 \times 10^3 \text{ W cm}^{-2}$ for 100 ns response time¹⁹, implying 3 – $40 \text{ pJ } \mu\text{m}^{-2}$. Hybrid devices, in which electrical feedback is incorporated to aid the nonlinearity, could also be used for optical circuits. Optical irradiances and switch energies can then be lower than the above figures but the extra electrical energy must also be properly included. Present values for an integrated MQW SEED device²⁰ are ~ 1 mW optical into a $120 \mu\text{m} \times 50 \mu\text{m}$ area with a $2 \mu\text{s}$ response time ($\sim 0.4 \text{ pJ } \mu\text{m}^{-2}$) plus electrical energy. It remains to be seen whether such highly complex devices can be scaled down to micrometre sizes. By comparison, the single nonlinear etalon devices can be easily pixellated on this scale with standard microfabrication techniques.

The typical switch contrast ratio for our NLIF devices in transmission is about $2:1$. Use of the devices in reflection gives NOR or NAND gate action—a necessity for computation—but often yields lower contrast (although higher gain). We have recently optimized the cavity reflectivities to allow for the increased level of on-resonance absorption and to thus improve this ratio; Fig. 3 shows a contrast ratio as high as $10:1$ and further improvement should be possible.

Digital optical circuits

Configuration for cascading of elements and for restoring optical logic. The technique for achieving gain, to ensure that the output of one element is sufficient to switch the succeeding one, we describe as 'hold and switch' (Fig. 1). By 'holding' the first element as close as possible to switching threshold, an arbitrarily small signal can be made to switch the device—limited only by the stability of the nonlinear element and its 'power supply', the holding laser beam. The change in output can thus be substantially larger than this (extra) switching signal. Thus, if

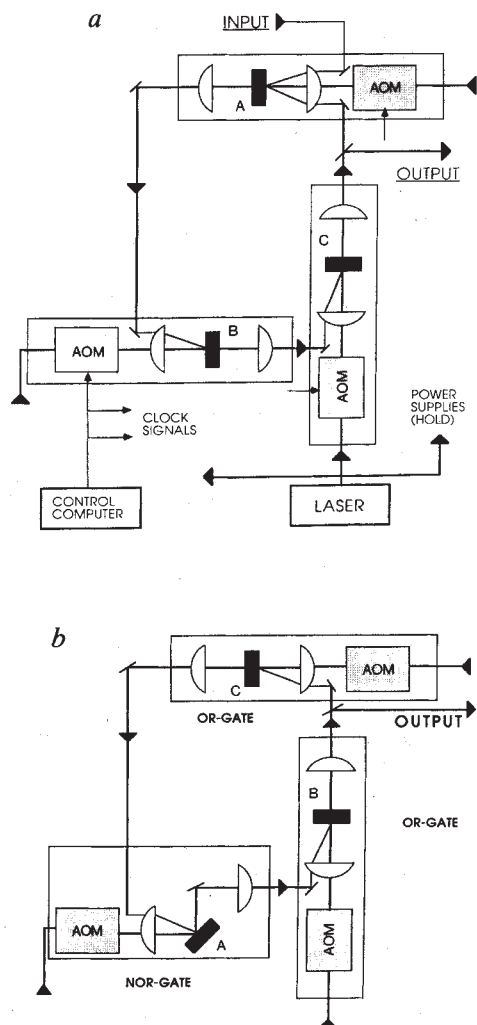


Fig. 4 *a*, Schematic diagram of three OR-gate loop circuit with external input. Each module contains one gate (labelled A, B or C); input/output optics and a computer controlled acousto-optic modulator (AOM) to clock the hold-power from the argon-ion laser. *b*, A similar loop circuit to *a*, but with one gate operating in reflection to provide a NOR-gate function.

this output is incident on a second device held equally close to threshold, it will clearly be enough to switch the second element and thus continue the circuit.

Practical limitations are that the power level of a typical argon-ion laser will fluctuate by ~3% and that drift of the NLIF switch power will preclude a closer hold over periods varying from minutes to hours, particularly, when operating dominantly 'on resonance' (where the internal optical field is high). In addition, it is desirable to 'over-switch' by a significant factor to reduce the effects of critical slowing down¹⁵. We typically hold to ~90% of threshold and over-switch by ~20%—determined by the output power change of the previous gate. The interconnection optics must preserve this situation for succeeding elements. We use both single and multilens imaging systems to collect and refocus the output signals onto the subsequent gates, as shown in Figs 4 and 7. Electronics normally separates power supplies from signal channels. We use our optical elements as three-port devices in a somewhat similar manner. 'Hold beams' (power supplies) are supplied to each element separately and the switching signals are introduced *off-axis* such that they are not collected by the down-line optical aperture (Fig. 1). Thus, provided the input signal is enough to switch the device, the output signal change is standard⁴. This gives restoring optical

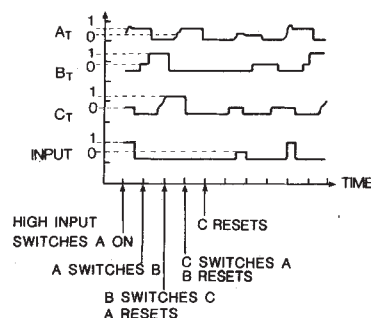


Fig. 5 Transmission output waveforms (A_T , B_T and C_T) from each gate in the circuit shown in Fig. 4*a*, plus input signal sequence. After transfer of logic-one around the loop all gates are reset. The gates then remain in their low state after a logic-zero is input—until the next logic-one.

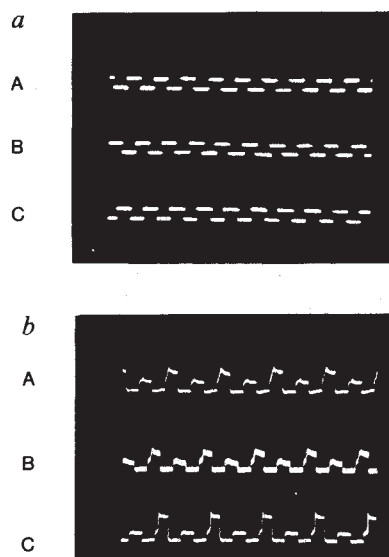


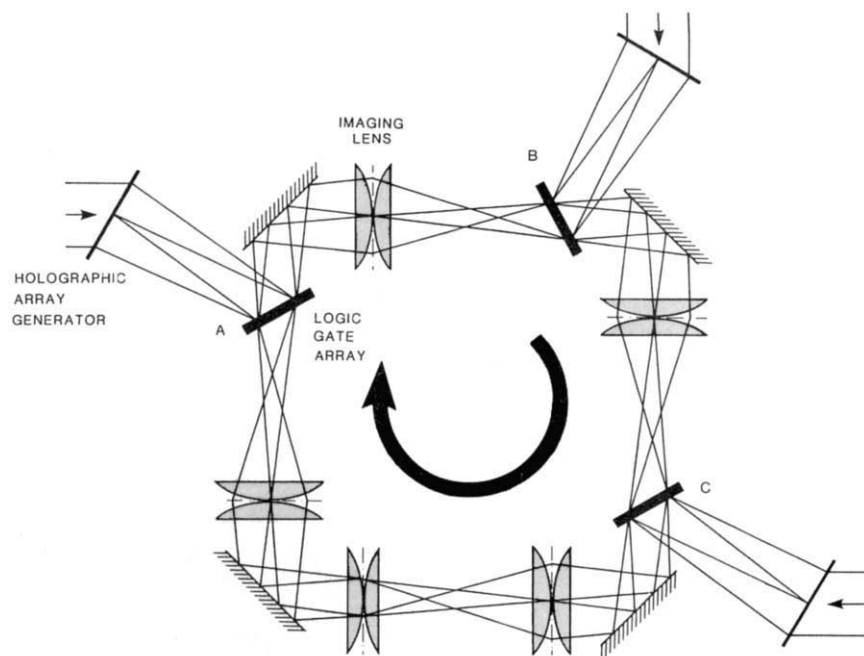
Fig. 6 Upper: the sequence of power levels of the three holding beams used in the inverting loop circuit shown in Fig. 4*b*. Lower: the resulting sequence of transmitted outputs obtained, showing inversion once every cycle.

logic and, hence, with sufficient independent holding beams, enables the circuit to be extended indefinitely.

'Lock and clock' architecture (buffered memory store). Figure 2 shows two forms of computer loop processors. These loop configurations are fundamental to many computer architectures. It is therefore essential when investigating the computing potential of any new type of digital gate to demonstrate its ability to pass data, without error, around such a loop. We have successfully constructed and operated several such circuits based on bistable NLIF devices.

To overcome the discrepancy between speed-of-light communication amongst optical gates and their slower switch rates, a 'lock and clock' architecture has been developed^{4,5}. This prevents the asynchronous transfer of data around the loop and thus avoids the consequent risk of a new signal appearing at the input to a gate before it has completed its response to the previous one.

The 'lock and clock' technique proceeds by stepping the holding beam on each gate between near switch point and zero, thus sequentially enabling and disabling gates. Thus one gate, when enabled, can respond to a transient input, latch (assuming it is held in a bistable region), pass on the datum to the next gate, which is brought from a disabled to enabled state, and



g. 7 Diagram of the configuration developed to monstrate an optical classical finite-state machine. Each hologram is used to split and focus the input beams so as to illuminate three gates on each of the LIF logic devices, A, B and C. The gate arrays are interconnected by direct imaging optics, as indicated. The holding power to each array, from a single argon-ion laser, is controlled by an acoustooptic modulator placed before the fan-out hologram.

finally reset as it is disabled. For a loop circuit containing three gates there is therefore a three-phase clock sequence for the holding beams in which no more than two gates are enabled at any one time. The clocking ensures that the datum is safely passed on to the next gate before each gate is reset. This sequence is the same as that for a three-element memory—the minimum required when just on-off holding beam control is used⁵. (With three-level holding beam control only two elements are required to store data in a loop⁴.)

Three OR-gate circuit with external input. Figure 4a shows schematically a successfully operated loop circuit based on off-axis addressed bistable NLIFs. Each element, when enabled by a standard holding beam, could be switched by a single input and thus act as an OR gate. The outputs were coupled to the next gates in the circuit. Gate A had two signal inputs, the previous gate output plus an external signal, corresponding to a total fan-in of three beams when the holding beam (power input) is included. The remaining gates acted as single-input OR gates.

The three elements were powered by an argon-ion laser via static attenuators and computer-controlled acoustooptic modulators (AOM), which, respectively, set the 'd.c.' levels and imposed the three-phase holding beam clock sequence. The external signals were determined by a fourth AOM.

Figure 5 shows the sequence of gate outputs (observed by detectors after each element). A valid logic-1 input was propagated around the loop in synchronization with the clock signal, followed, on the next loop cycle, by a signal too small to switch (valid logic-zero), which leaves the three gates in their logic-zero states. (The loop was reset to zero before each full cycle.) This transmission-coupled circuit has thus shown that an extensible system of restoring optical logic can be constructed with all-optical bistable devices.

Inverting OR/NOR-gate circuit. Figure 4b shows a second loop circuit, similar to the previous one but with one element (A) operated in a reflection-output mode. This acts as a NOR-gate so that when taken with two OR-gates a loop circuit with overall inversion is created. Because such a circuit automatically inverts the circulating datum on each cycle, no external input is needed to test for proper operation.

Figure 6 shows both the three-phase clock sequence applied to the three gates and the detected outputs from each gate. The transmission (OR) output was monitored. On reaching gate A,

logic-one is rendered logic-zero on one cycle and zero becomes one on the succeeding cycle. Both restoring logic and the inversion function are convincingly demonstrated.

Parallel circuits: an optical classical-finite-state machine. We are constructing circuits similar to those described above, but extended to show parallel transfer of data around a logic loop. Figure 7 shows, schematically, a system of three, parallel inverting, NOR-gate loop circuits. This is an all-reflection coupled system in which the higher absolute power change obtained from the reflected outputs is exploited. Parallelism is achieved by illuminating an array of points on each (uniform) bistable element using separate holding beams. Each gate, thus defined, acts as an independent switch.

To generate the pattern of holding beams we have developed holographic optical element (HOE) arrays, designed to produce a corresponding set of focal spots when illuminated by a single, uniform irradiance, parallel beam. These HOEs have been developed in-house with dichromated gelatin (DCG) as the recording medium and they yield high diffraction efficiencies of approximately 90%. Rectangular lenslet arrays containing up to 100 elements have been successfully produced. For the system in Fig. 7, a simple, triangular, pattern of three lenslets has been used to demonstrate a minimum level of parallelism. Because the signal connections are achieved by direct imaging of one bistable NLIF onto the next, the number of parallel channels can be increased simply by substituting the hologram with one that generates a larger number of focused beamlets.

The gates themselves are operated as BEAT devices¹⁷, in which the transmitted output is absorbed in a thin opaque film on the outer surface of the filter structure. In addition to some increase in sensitivity, this technique considerably eases the problem of efficiently coupling the signal beams into the gate by providing an angle insensitive, full absorber as the input port.

The circuit is clocked, in the usual three-phase sequence, by AOM control of each of the three power beams before they are fanned-out by the HOE array. Switching of the array of bistable all-optical gates by the outputs from the preceding array has been successfully demonstrated.

By comparing Fig. 7 with Fig. 2a it can be seen that this circuit has many of the features of a classical finite-state machine. Two of the gate arrays can be regarded as a memory, into and from which data can be clocked, and the third array as a (clocked) logic unit. In this case a simple logical inversion is

applied to the data before directly recycling. But it is important that such an optical, classical finite-state machine is demonstrated to prove the feasibility of parallel-optical digital logic.

Conclusions

The operation of a number of digital all-optical loop circuits, based on bistable nonlinear interference filter (NLIF) devices, has proven the capability of all-optical switches of this type to form systems of extensible optical restoring logic. Furthermore the successful switching of cascaded arrays of such all-optical digital gates has confirmed their anticipated potential for creating parallel processing systems.

We have chosen to use electronic methods for control and clocking of these circuits. There is nothing 'optically impure' in this approach in the régime where switch times are greater than total-loop time. Indeed, it provides an important interface with conventional electronics. For various logic decision purposes one could anticipate that the 'optical classical finite-state-machine', as depicted, could be added to an electronic machine to multiply the processor power by the number of parallel optical channels (single instruction multiple data-stream).

The interference film devices that have been used lend themselves to parallelism because they are intrinsically uniform to a high degree. The pass-band (~ 4 nm FWHM) has been measured to be constant in width and centre wavelength to $\leq 3\%$ over an

area of ~ 25 mm² with a spatially mapping Fourier transform interferometric spectrometer.

The effects of thermal diffusion on spot-size scaling, gate cross-talk and switching speeds have been studied both theoretically and experimentally^{11,14}. Sample pixellation is essential to prevent cross-talk in large arrays. We are pursuing fabrication techniques that include pixellation of the active layers and of thermally designed substrates. Even at the present switching power levels, 1 mW for 10 μ m spot sizes and 30 μ s switching times, the achievable parallelism is 10^3 gates per watt. For pixellated thermal devices optimized for low-energy switching the predicted parameters are 1 pJ μ m⁻² (for example, 100 μ W, 10 μ m², 100 ns)^{14,21}. This would yield up to 10^{11} Hz W⁻¹. Given easily achievable heat sinking of 10 W cm⁻² this implies a switching capacity of 10^{12} Hz cm⁻². Such performance would permit experimental assessment of a wide variety of all-optical parallel information processing in the near future.

We acknowledge the essential contributions of J. G. H. Mathew, M. R. Taghizadeh, N. Craft, I. Redmond and R. J. Campbell to the experimental work. Partially supported by the UK Science and Engineering Research Council, Joint Opto-Electronic Research Scheme (JOERS); the work was carried out within the framework of the European Joint Optical Bistability Project (EJOB) of the Commission of the European Communities.

Received 27 August; accepted 10 November 1986.

- Walker, A. C., Tooley, F. A. P., Prise, M. E., Mathew, J. G. H., Kar, A. K., Taghizadeh, M. R. & Smith, S. D. *Phil. Trans. R. Soc. A* **313**, 249-256 (1984).
- Smith, S. D., Mathew, J. G. H., Taghizadeh, M. R., Walker, A. C., Wherrett, B. S. & Hendry, A. *Optics Commun.* **51**, 357-362 (1984).
- Smith, S. D. *Nature* **316**, 319-324 (1985).
- Walker, A. C. *Appl. Opt.* **25**, 1575-1585 (1986).
- Wherrett, B. S. *Appl. Opt.* **24**, 2876-2883 (1985).
- Smith, S. D. *et al. Opt. Engng* **24**, 569-574 (1985).
- Mead, C. & Conway, C. *An Introduction to VLSI Systems* (Addison-Wesley, Reading, Massachusetts, 1980).
- Goodman, J. W. *Proc. Inst. Radio Engrs* **65**, 29-32 (1977).
- Lohman, A. W. *Optical Bistability III* (eds Gibbs, H. M., Smith, S. D., Mandel, P. & Peyghambarian, N.) 2-3 (Springer, Berlin, 1986).
- Huang, A. *Proc. Inst. Radio Engrs* **72**, 780-786 (1984).
- Abraham, E. & Halley, J. M. *Appl. Phys. A* (in the press).
- Bigot, J. Y., Daunois, A., Leonelli, R., Sence, M., Mathew, J. G. H., Smith, S. D. & Walker, A. C. *Appl. Phys. Lett.* **49**, 844-846 (1986).

- Wherrett, B. S., Hutchings, D. & Russell, D. J. *opt. Soc. Am. B* **3**, 351-362 (1986).
- Janossy, I., Mathew, J. G. H., Abraham, E., Taghizadeh, M. R. & Smith, S. D. *IEEE J. Quantum Electron.* (in the press).
- Mathew, J. G. H., Taghizadeh, M. R., Abraham, E., Janossy, I. & Smith, S. D. *Optical Bistability III* (eds Gibbs, H. M., Smith, S. D., Mandel, P. & Peyghambarian, N.) 57-60 (Springer, Berlin, 1986).
- Janossy, I., Taghizadeh, M. R., Mathew, J. G. H. & Smith, S. D. *IEEE J. Quantum Electron.* **21**, 1447-1452 (1985).
- Walker, A. C. *Optics Commun.* **59**, 145-150 (1986).
- Gibbs, H. M., Tarnag, S. S., Jewell, J. L., Weinberger, D. A. & Tai, K. *Appl. Phys. Lett.* **41**, 221-223 (1982).
- Ovadia, S., Gibbs, H. M., Jewell, J. L., Sarid, D. & Peyghambarian, N. *Opt. Engng* **24**, 565-568 (1985).
- Miller, D. A. B., Henry, J. E., Gossard, A. C. & English, J. H. *Appl. Phys. Lett.* **49**, 821-823 (1986).
- Wherrett, B. S. in *From Optical Bistability Towards Optical Computing* (Proc. European Joint Optical Bistability Project) (eds Smith, S. D., Mandel, P. & Wherrett, B. S.) Ch. 5.7 (North Holland, Amsterdam, in the press).

Mitochondrial DNA and human evolution

Rebecca L. Cann*, Mark Stoneking & Allan C. Wilson

Department of Biochemistry, University of California, Berkeley, California 94720, USA

Mitochondrial DNAs from 147 people, drawn from five geographic populations have been analysed by restriction mapping. All these mitochondrial DNAs stem from one woman who is postulated to have lived about 200,000 years ago, probably in Africa. All the populations examined except the African population have multiple origins, implying that each area was colonised repeatedly.

MOLECULAR biology is now a major source of quantitative and objective information about the evolutionary history of the human species. It has provided new insights into our genetic divergence from apes¹⁻⁸ and into the way in which humans are related to one another genetically⁹⁻¹⁴. Our picture of genetic evolution within the human species is clouded, however, because it is based mainly on comparisons of genes in the nucleus. Mutations accumulate slowly in nuclear genes. In addition, nuclear genes are inherited from both parents and mix in every generation. This mixing obscures the history of individuals and

allows recombination to occur. Recombination makes it hard to trace the history of particular segments of DNA unless tightly linked sites within them are considered.

Our world-wide survey of mitochondrial DNA (mtDNA) adds to knowledge of the history of the human gene pool in three ways. First, mtDNA gives a magnified view of the diversity present in the human gene pool, because mutations accumulate in this DNA several times faster than in the nucleus¹⁵. Second, because mtDNA is inherited maternally and does not recombine¹⁶, it is a tool for relating individuals to one another. Third, there are about 10^{16} mtDNA molecules within a typical human and they are usually identical to one another¹⁷⁻¹⁹. Typical mam-

* Present address: Department of Genetics, University of Hawaii, Honolulu, Hawaii 96822.

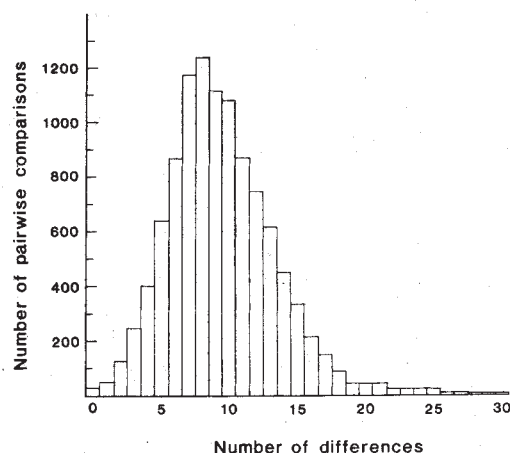


Fig. 1 Histogram showing the number of site differences between restriction maps of mtDNA for all possible pairs of 147 human beings.

malian females consequently behave as haploids, owing to a bottleneck in the genetically effective size of the population of mtDNA molecules within each oocyte²⁰. This maternal and haploid inheritance means that mtDNA is more sensitive than nuclear DNA to severe reductions in the number of individuals in a population of organisms¹⁵. A pair of breeding individuals can transmit only one type of mtDNA but carry four haploid sets of nuclear genes, all of which are transmissible to offspring. The fast evolution and peculiar mode of inheritance of mtDNA provide new perspectives on how, where and when the human gene pool arose and grew.

Restriction maps

MtDNA was highly purified from 145 placentas and two cell lines, HeLa and GM 3043, derived from a Black American and an aboriginal South African (IKung), respectively. Most placentas (98) were obtained from US hospitals, the remainder coming from Australia and New Guinea. In the sample, there were representatives of 5 geographic regions: 20 Africans (representing the sub-Saharan region), 34 Asians (originating from China, Vietnam, Laos, the Philippines, Indonesia and Tonga), 46 Caucasians (originating from Europe, North Africa, and the Middle East), 21 aboriginal Australians, and 26 aboriginal New Guineans. Only two of the 20 Africans in our sample, those bearing mtDNA types 1 and 81 (see below) were born in sub-Saharan Africa. The other 18 people in this sample are Black Americans, who bear many non-African nuclear genes probably contributed mainly by Caucasian males. Those males would not be expected to have introduced any mtDNA to the Black American population. Consistent with our view that most of

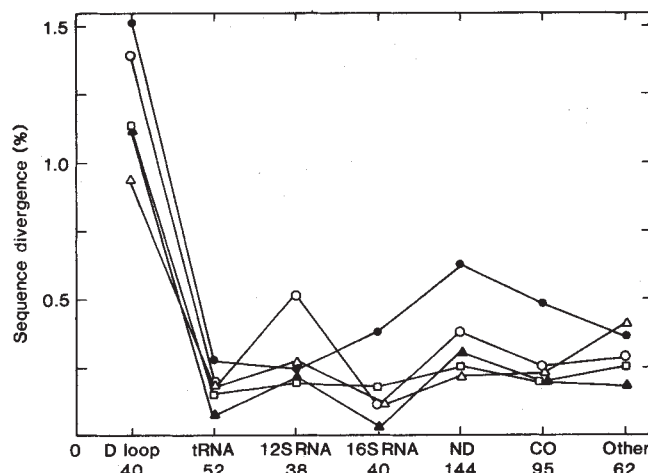


Fig. 2 Sequence divergence within 5 geographic areas for each of 7 functional regions in human mtDNA. Sequence divergence (δ_x) was estimated from comparisons of restriction maps²⁶. Symbols for the 5 races are: ●, Africa; ○, Asia; △, Australia; □, Caucasian; ▲, New Guinea. Along the horizontal axis are the numbers of restriction sites in each functional region (D loop, transfer RNA genes, 12S and 16S ribosomal RNA genes, NADH dehydrogenase subunits 1–5, cytochrome oxidase subunits and other protein-coding regions).

these 18 people are a reliable source of African mtDNA, we found that 12 of them bear restriction site markers known²¹ to occur exclusively or predominantly in native sub-Saharan Africans (but not in Europeans, Asians or American Indians nor, indeed, in all such Africans). The mtDNA types in these 12 people are 2–7, 37–41 and 82 (see below). Methods used to purify mtDNA and more detailed ethnographic information on the first four groups are as described^{17,22}; the New Guineans are mainly from the Eastern Highlands of Papua New Guinea²³.

Each purified mtDNA was subjected to high resolution mapping^{22–24} with 12 restriction enzymes (*Hpa*I, *Ava*II, *Fnu*DII, *Hha*I, *Hpa*II, *Mbo*I, *Taq*I, *Rsa*I, *Hinf*I, *Hae*III, *Aha*I and *Dde*I). Restriction sites were mapped by comparing observed fragment patterns to those expected from the known human mtDNA sequence²⁵. In this way, we identified 467 independent sites, of which 195 were polymorphic (that is, absent in at least one individual). An average of 370 restriction sites per individual were surveyed, representing about 9% of the 16,569 base-pair human mtDNA genome.

Map comparisons

The 147 mtDNAs mapped were divisible into 133 distinct types. Seven of these types were found in more than one individual;

Table 1 MtDNA divergence within and between 5 human populations

Population	% sequence divergence				
	1	2	3	4	5
1. African	0.47	0.04	0.04	0.05	0.06
2. Asian	0.45	0.35	0.01	0.02	0.04
3. Australian	0.40	0.31	0.25	0.03	0.04
4. Caucasian	0.40	0.31	0.27	0.23	0.05
5. New Guinean	0.42	0.34	0.29	0.29	0.25

The divergence is calculated by a published method²⁶. Values of the mean pairwise divergence between individuals within populations (δ_x) appear on the diagonal. Values below the diagonal (δ_{xy}) are the mean pairwise divergences between individuals belonging to two different populations, X and Y. Values above the diagonal (δ) are interpopulation divergences, corrected for variation within those populations with equation (1).

Table 2 Clusters of mtDNA types that are specific to one geographic region

Geographic region	Number of region-specific clusters	Mean pairwise distance within clusters*	Average age of clusters†
Africa	1‡	0.36	90–180
Asia	27	0.21	53–105
Australia	15	0.17	43–85
Europe	36	0.09	23–45
New Guinea	7	0.11	28–55

* For clusters represented by two or more individuals (and calculated for individuals, not for mtDNA types) in Fig. 3.

† Average age in thousands of years based on the assumption that mtDNA divergence occurs at the rate of 2–4% per million years^{15,30}.

‡ Assuming that Africa is the source, there is only one African cluster.

no individual contained more than one type. None of the seven shared types occurred in more than one of the five geographic regions. One type, for example, was found in two Australians. Among Caucasians, another type occurred three times and two more types occurred twice. In New Guinea, two additional types were found three times and the seventh case involved a type found in six individuals.

A histogram showing the number of restriction site differences between pairs of individuals is given in Fig. 1; the average number of differences observed between any two humans is 9.5. The distribution is approximately normal, with an excess of pairwise comparisons involving large numbers of differences.

From the number of restriction site differences, we estimated the extent of nucleotide sequence divergence²⁶ for each pair of individuals. These estimates ranged from zero to 1.3 substitutions per 100 base pairs, with an average sequence divergence of 0.32%, which agrees with that of Brown¹⁷, who examined only 21 humans.

Table 1 gives three measures of sequence divergence within and between each of the five populations examined. These measures are related to one another by equation (1):

$$\delta = \delta_{xy} - 0.5(\delta_x + \delta_y) \quad (1)$$

where δ_x is the mean pairwise divergence (in percent) between individuals within a single population (X), δ_y is the corresponding value for another population (Y), δ_{xy} is the mean pairwise divergence between individuals belonging to two different populations (X and Y), and δ is a measure of the interpopulation divergence corrected for intrapopulation divergence. Africans as a group are more variable ($\delta_x = 0.47$) than other groups. Indeed, the variation within the African population is as great as that between Africans and any other group ($\delta_{xy} = 0.40$ – 0.45). The within-group variation of Asians ($\delta_x = 0.35$) is also comparable to that which exists between groups. For Australians, Caucasians, and New Guineans, who show nearly identical amounts of within-group variation ($\delta_x = 0.23$ – 0.25), the variation between groups slightly exceeds that within groups.

When the interpopulation distances (δ_{xy}) are corrected for intrapopulation variation (Table 1), they become very small ($\delta = 0.01$ – 0.06). The mean value of the corrected distance among populations ($\delta = 0.04$) is less than one-seventh of the mean distance between individuals within a population (0.30). Most of the mtDNA variation in the human species is therefore shared between populations. A more detailed analysis supports this view²⁷.

Functional constraints

Figure 2 shows the sequence divergence (δ_x) calculated for each population across seven functionally distinct regions of the mtDNA genome. As has been found before^{24,27,28}, the most variable region is the displacement loop ($\delta_x = 1.3$), the major noncoding portion of the mtDNA molecule, and the least variable region is the 16S ribosomal RNA gene ($\delta_x = 0.2$). In general, Africans are the most diverse and Asians the next most, across all functional regions.

Evolutionary tree

A tree relating the 133 types of human mtDNA and the reference sequence (Fig. 3) was built by the parsimony method. To interpret this tree, we make two assumptions, both of which have extensive empirical support: (1) a strictly maternal mode of mtDNA transmission (so that any variant appearing in a group of lineages must be due to a mutation occurring in the ancestral lineage and not recombination between maternal and paternal genomes) and (2) each individual is homogeneous for its multiple mtDNA genomes. We can therefore view the tree as a genealogy linking maternal lineages in modern human populations to a common ancestral female (bearing mtDNA type a).

Many trees of minimal or near-minimal length can be made from the data; all trees that we have examined share the

following features with Fig. 3. (1) two primary branches, one composed entirely of Africans, the other including all 5 of the populations studied; and (2) each population stems from multiple lineages connected to the tree at widely dispersed positions. Since submission of this manuscript, Horai *et al.*²⁹ built a tree for our samples of African and Caucasian populations and their sample of a Japanese population by another method; their tree shares these two features.

Among the trees investigated was one consisting of five primary branches with each branch leading exclusively to one of the five populations. This tree, which we call the population-specific tree, requires 51 more point mutations than does the tree of minimum length in Fig. 3. The minimum-length tree requires fewer changes at 22 of the 93 phylogenetically-informative restriction sites than does the population-specific tree, while the latter tree required fewer changes at four sites; both trees require the same number of changes at the remaining 67 sites. The minimum-length tree is thus favoured by a score of 22 to 4. The hypothesis that the two trees are equally compatible with the data is statistically rejected, since 22:4 is significantly different from the expected 13:13. The minimum-length tree is thus significantly more parsimonious than the population-specific tree.

African origin

We infer from the tree of minimum length (Fig. 3) that Africa is a likely source of the human mitochondrial gene pool. This inference comes from the observation that one of the two primary branches leads exclusively to African mtDNAs (types 1–7, Fig. 3) while the second primary branch also leads to African mtDNAs (types 37–41, 45, 46, 70, 72, 81, 82, 111 and 113). By postulating that the common ancestral mtDNA (type a in Fig. 3) was African, we minimize the number of intercontinental migrations needed to account for the geographic distribution of mtDNA types. It follows that b is a likely common ancestor of all non-African and many African mtDNAs (types 8–134 in Fig. 3).

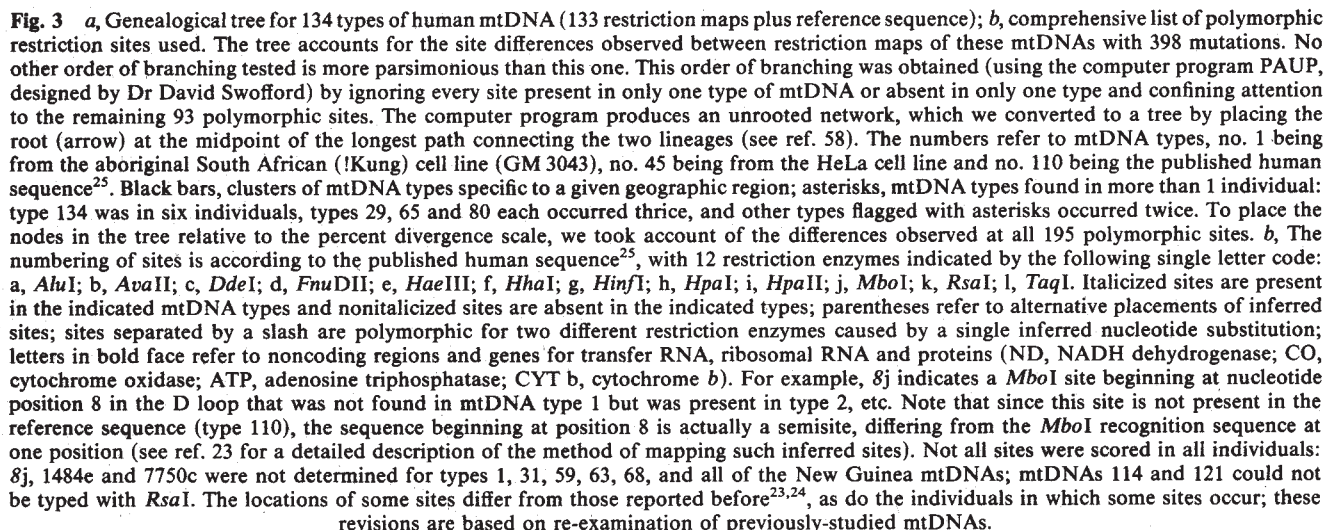
Multiple lineages per race

The second implication of the tree (Fig. 3)—that each non-African population has multiple origins—can be illustrated most simply with the New Guineans. Take, as an example, mtDNA type 49, a lineage whose nearest relative is not in New Guinea, but in Asia (type 50). Asian lineage 50 is closer genealogically to this New Guinea lineage than to other Asian mtDNA lineages. Six other lineages lead exclusively to New Guinean mtDNAs, each originating at a different place in the tree (types 12, 13, 26–29, 65, 95 and 127–134 in Fig. 3). This small region of New Guinea (mainly the Eastern Highlands Province) thus seems to have been colonised by at least seven maternal lineages (Tables 2 and 3).

In the same way, we calculate the minimum numbers of female lineages that colonised Australia, Asia and Europe (Tables 2 and 3). Each estimate is based on the number of region-specific clusters in the tree (Fig. 3, Tables 2 and 3). These numbers, ranging from 15 to 36 (Tables 2 and 3), will probably rise as more types of human mtDNA are discovered.

Tentative time scale

A time scale can be affixed to the tree in Fig. 3 by assuming that mtDNA sequence divergence accumulates at a constant rate in humans. One way of estimating this rate is to consider the extent of differentiation within clusters specific to New Guinea (Table 2; see also refs 23 and 30), Australia³⁰ and the New World³¹. People colonised these regions relatively recently: a minimum of 30,000 years ago for New Guinea³², 40,000 years ago for Australia³³, and 12,000 years ago for the New World³⁴. These times enable us to calculate that the mean rate of mtDNA divergence within humans lies between two and four percent per million years; a detailed account of this calculation appears



When did the migrations from Africa take place? The oldest of the clusters of mtDNA types to contain no African members stems from ancestor *c* and included types 11–29 (Fig. 3). The

Two previous studies of human mtDNA have included African individuals^{21,28}, both support an African origin for the human mtDNA gene pool. Johnson *et al.*²¹ surveyed ~40 restriction sites in each of 200 mtDNAs from Africa, Asia, Europe and the New World, and found 35 mtDNA types. This much smaller number of mtDNA types probably reflects the inability of their methods to distinguish between mtDNAs that differ by less than 0.3% and may account for the greater clustering of mtDNA

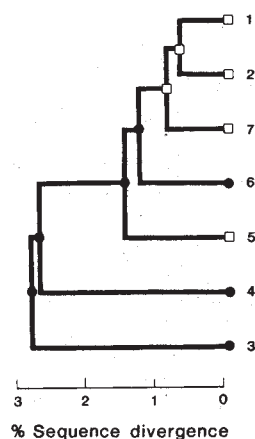


Fig. 4 Genealogical tree relating the nucleotide sequences of D loops from seven human mtDNAs. This tree, which requires fewer mutations than any other branching order, was constructed by the PAUP computer program from the 900-bp sequences determined by Greenberg *et al.*²⁸ and was rooted by the midpoint method. Symbols: ●, African origin (Black American); □, Caucasian.

types by geographic origin that they observed. (By contrast, our methods distinguish between mtDNAs that differ by 0.03%.) Although Johnson *et al.* favoured an Asian origin, they too found that Africans possess the greatest amount of mtDNA variability and that a midpoint rooting of their tree leads to an African origin.

Greenberg *et al.*²⁸ sequenced the large noncoding region, which includes the displacement loop (D loop), from four Caucasians and three Black Americans. A parsimony tree for these seven D loop sequences, rooted by the midpoint method, appears in Fig. 4. This tree indicates (1) a high evolutionary rate for the D loop (at least five times faster than other mtDNA regions), (2) a greater diversity among Black American D loop sequences, and (3) that the common ancestor was African.

Nuclear DNA studies

Estimates of genetic distance based on comparative studies of nuclear genes and their products differ in kind from mtDNA estimates. The latter are based on the actual number of mutational differences between mtDNA genomes, while the former rely on differences in the frequencies of molecular variants measured between and within populations. Gene frequencies can be influenced by recombination, genetic drift, selection, and migration, so the direct relationship found between time and mutational distance for mtDNA would not be expected for genetic distances based on nuclear DNA. But studies based on polymorphic blood groups, red cell enzymes, and serum proteins show that (1) differences between racial groups are smaller than those within such groups and (2) the largest gene frequency

differences are between Africans and other populations, suggesting an African origin for the human nuclear gene pool^{11,12,35}. More recent studies of restriction site polymorphisms in nuclear DNA^{14,36-42} support these conclusions.

Relation to fossil record

Our tentative interpretation of the tree (Fig. 3) and the associated time scale (Table 3) fits with one view of the fossil record: that the transformation of archaic to anatomically modern forms of *Homo sapiens* occurred first in Africa⁴³⁻⁴⁵, about 100,000–140,000 years ago, and that all present-day humans are descendants of that African population. Archaeologists have observed that blades were in common use in Africa 80–90 thousand years ago, long before they replaced flake tools in Asia or Europe^{46,47}. But the agreement between our molecular view and the evidence from palaeoanthropology and archaeology should be treated cautiously for two reasons. First, there is much uncertainty about the ages of these remains. Second, our placement of the common ancestor of all human mtDNA diversity in Africa 140,000–280,000 years ago need not imply that the transformation to anatomically modern *Homo sapiens* occurred in Africa at this time. The mtDNA data tell us nothing of the contributions to this transformation by the genetic and cultural traits of males and females whose mtDNA became extinct.

An alternative view of human evolution rests on evidence that *Homo* has been present in Asia as well as in Africa for at least one million years⁴⁸ and holds that the transformation of archaic to anatomically modern humans occurred in parallel in different parts of the Old World^{33,49}. This hypothesis leads us to expect genetic differences of great antiquity within widely separated parts of the modern pool of mtDNAs. It is hard to reconcile the mtDNA results with this hypothesis. The greatest divergences within clusters specific to non-African parts of the World correspond to times of only 90,000–180,000 years. This might imply that the early Asian *Homo* (such as Java man and Peking man) contributed no surviving mtDNA lineages to the gene pool of our species. Consistent with this implication are features, found recently in the skeletons of the ancient Asian forms, that make it unlikely that Asian *erectus* was ancestral to *Homo sapiens*⁵⁰⁻⁵². Perhaps the non-African *erectus* population was replaced by *sapiens* migrants from Africa; incomplete fossils indicating the possible presence of early modern humans in western Asia at Zuttiyeh (75,000–150,000 years ago) and Qafzeh (50,000–70,000 years ago) might reflect these first migrations^{45,53}.

If there was hybridization between the resident archaic forms in Asia and anatomically modern forms emerging from Africa, we should expect to find extremely divergent types of mtDNA in present-day Asians, more divergent than any mtDNA found in Africa. There is no evidence for these types of mtDNA among the Asians studied^{21,54-56}. Although such archaic types of mtDNA could have been lost from the hybridizing population, the probability of mtDNA lineages becoming extinct in an expanding population is low⁵⁷. Thus we propose that *Homo*

Table 3 Ancestors, lineages and extents of divergence in the genealogical tree for 134 types of human mtDNA

Ancestor	Total	No. of descendant lineages or clusters specific to a region					% divergence	Age*
		Africa	Asia	Australia	Europe	N. Guinea		
a	7	1	0	0	0	0	0.57	143–285
b	2	0	1	0	0	0	0.45	112–225
c	20	0	7	3	1	3	0.43	108–215
d	2	0	0	1	1	0	0.39	98–195
e	14	2	2	4	2	0	0.34	85–170
f	19	1	7	4	4	1	0.30	75–150
g	10	2	3	2	2	1	0.28	70–140
h	30	2	4	0	15	1	0.27	68–135
i	8	1	0	0	6	0	0.26	65–130
j	22	1	3	1	5	1	0.25	62–125
All	134	10	27	15	36	7	—	—

* Assuming that the mtDNA divergence rate is 2–4% per million years^{15,30}.

erectus in Asia was replaced without much mixing with the invading *Homo sapiens* from Africa.

Conclusions and prospects

Studies of mtDNA suggest a view of how, where and when modern humans arose that fits with one interpretation of evidence from ancient human bones and tools. More extensive molecular comparisons are needed to improve our rooting of the mtDNA tree and the calibration of the rate of mtDNA divergence within the human species. This may provide a more reliable time scale for the spread of human populations and better estimates of the number of maternal lineages involved in founding the non-African populations.

It is also important to obtain more quantitative estimates of

the overall extent of nuclear DNA diversity in both human and African ape populations. By comparing the nuclear and mitochondrial DNA diversities, it may be possible to find out whether a transient or prolonged bottleneck in population size accompanied the origin of our species¹⁵. Then a fuller interaction between palaeoanthropology, archaeology and molecular biology will allow a deeper analysis of how our species arose.

We thank the Foundation for Research into the Origin of Man, the National Science Foundation and the NIH for support. We also thank P. Andrews, K. Bhatia, F. C. Howell, W. W. Howells, R. L. Kirk, E. Mayr, E. M. Prager, V. M. Sarich, C. Stringer and T. White for discussion and help in obtaining placentas.

Received 17 March; accepted 7 November 1986.

- Goodman, M. *Hum. Biol.* **35**, 377-424 (1963).
- Sarich, V. M. & Wilson, A. C. *Science* **158**, 1200-1203 (1967).
- King, M. C. & Wilson, A. C. *Science* **188**, 107-116 (1975).
- Ferris, S. D., Wilson, A. C. & Brown, W. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2432-2436 (1981).
- Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. *J. molec. Evol.* **18**, 225-239 (1982).
- Sibley, C. G. & Ahlquist, J. E. *J. molec. Evol.* **20**, 2-15 (1984).
- Bianchi, N. O., Bianchi, M. S., Cleaver, J. E. & Wolf, S. J. *J. molec. Evol.* **22**, 323-333 (1985).
- O'Brien, S. J. *et al. Nature* **317**, 140-144 (1985).
- Cavalli-Sforza, L. L. *Proc. R. Soc. Lond. B164*, 362-379 (1966).
- Cavalli-Sforza, L. L. *Sci. Am.* **231**(3), 81-89 (1974).
- Nei, M. & Roychoudhury, A. K. *Evol. Biol.* **14**, 1-59 (1982).
- Nei, M. in *Population Genetics and Molecular Evolution* (eds Ohta, T. & Aoki, K.) 41-64 (Japan Sci. Soc. Press, Tokyo, 1985).
- Constans, J. *et al. Am. J. phys. Anthropol.* **68**, 107-122 (1985).
- Wainscoat, J. S. *et al. Nature* **319**, 491-493 (1986).
- Wilson, A. C. *et al. Biol. J. Linn. Soc.* **26**, 375-400 (1985).
- Olivo, P. D., Van de Walle, M. J., Laipis, P. J. & Hauswirth, W. W. *Nature* **306**, 400-402 (1983).
- Brown, W. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3605-3609 (1980).
- Monnat, R. J. & Loeb, L. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2895-2899 (1985).
- Monnat, R. J., Maxwell, C. L. & Loeb, L. A. *Cancer Res.* **45**, 1809-1814 (1985).
- Hauswirth, W. W. & Laipis, P. J. in *Achievements and Perspectives in Mitochondrial Research, Vol. 2 Biogenesis* (eds E. Quagliariello, E. C. Slater, F. Palmieri, C. Saccone & A. M. Kroon) 49-60 (Elsevier, New York, 1986).
- Johnson, M. J. *et al. J. molec. Evol.* **19**, 255-271 (1983).
- Cann, R. L. Ph.D. Thesis, Univ. California, Berkeley, California (1982).
- Stoneking, M., Bhatia, K. & Wilson, A. C. in *Genetic Variation and Its Maintenance in Tropical Populations* (eds Roberts, C. F. & Destefano, G.) 87-100 (Cambridge University Press, Cambridge, 1986).
- Cann, R. L., Brown, W. M. & Wilson, A. C. *Genetics* **106**, 479-499 (1984).
- Anderson, S. *et al. Nature* **290**, 457-465 (1981).
- Nei, M. & Tajima, F. *Genetics* **105**, 207-217 (1983).
- Whitman, T. S., Clark, A. G., Stoneking, M., Cann, R. L. & Wilson, A. C. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Greenberg, B. D., Newbold, J. E. & Sugino, A. *Gene* **21**, 33-49 (1983).

- Horai, S., Gojobori, T. & Matsunaga, E. *Jap. J. Genet.* **61**, 271-275 (1986).
- Stoneking, M., Bhatia, K. & Wilson, A. C. *Cold Spring Harb. Symp. quant. Biol.* **51**, 433-439 (1986).
- Wallace, D. C., Garrison, K. & Knowler, W. C. *Am. J. phys. Anthropol.* **68**, 149-155 (1985).
- Jones, R. A. *Rev. Anthropol.* **8**, 445-466 (1979).
- Wolpoff, M. H., Wu, X. Z. & Thorne, A. G. in *Origins of Modern Humans: A World Survey of the Fossil Evidence* (eds Smith, F. H. & Spencer, F.) 411-484 (Liss, New York, 1984).
- Owen, R. C. in *Origins of Modern Humans: A World Survey of the Fossil Evidence* (eds Smith, F. H. & Spencer, F.) 517-564 (Liss, New York, 1984).
- Mourant, A. E. *et al. The Distribution of the Human Blood Groups and Other Polymorphism* (Oxford University Press, Oxford, 1978).
- Murray, J. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3486-3490 (1984).
- Cooper, D. N. & Schmidtke, J. *Hum. Genet.* **66**, 1-16 (1984).
- Cooper, D. N. *et al. Hum. Genet.* **69**, 201-205 (1985).
- Hill, A. V. S., Nicholls, R. D., Thein, S. L. & Higgs, D. R. *Cell* **42**, 809-819 (1985).
- Chapman, B. S., Vincent, K. A. & Wilson, A. C. *Genetics* **112**, 79-92 (1986).
- Chakravarti, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 6085-6089 (1984).
- Chakravarti, A. *et al. Am. J. hum. Genet.* **36**, 1239-1258 (1984).
- Rightmire, G. P. in *Origins of Modern Humans: A World Survey of the Fossil Evidence* (eds Smith, F. H. & Spencer, F.) 295-326 (Liss, New York, 1984).
- Bräuer, G. *Courier Forsch. Inst. Senckenberg* **69**, 145-165 (1984).
- Bräuer, G. in *Origins of Modern Humans: A World Survey of the Fossil Evidence* (eds Smith, F. H. & Spencer, F.) 327-410 (Liss, New York, 1984).
- Isaac, G. *Phil. Trans. R. Soc. Lond. B292*, 177-188 (1981).
- Clark, J. D. *Proc. Br. Acad. Lond.* **67**, 163-192 (1981).
- Pope, G. G. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4988-4992 (1983).
- Coon, C. S. *The Origin of Races* (Knopf, New York, 1962).
- Stringer, C. B. *Courier Forsch. Inst. Senckenberg* **69**, 131-143 (1984).
- Andrews, P. *Courier Forsch. Inst. Senckenberg* **69**, 167-175 (1984).
- Andrew, P. *New Scientist* **102**, 24-26 (1984).
- Stringer, C. B., Hublin, J. J. & Vandermeersch, B. in *Origins of Modern Humans: A World Survey of the Fossil Evidence* (eds Smith, F. H. & Spencer, F.) 51-136 (Liss, New York, 1984).
- Horai, S., Gojobori, T. & Matsunaga, E. *Hum. Genet.* **68**, 324-332 (1984).
- Bonne-Tamir, B. *et al. Am. J. hum. Genet.* **38**, 341-351 (1986).
- Horai, S. & Matsunaga, E. *Hum. Genet.* **72**, 105-117 (1986).
- Avise, J. C., Neigel, J. E. & Arnold, J. J. *molec. Evol.* **20**, 99-105 (1984).
- Farris, J. S. *Am. Nat.* **106**, 645-668 (1972).

LETTERS TO NATURE

Superluminal motion in the double-lobed quasar 3C263

J. A. Zensus*, D. H. Hough*† & R. W. Porcas‡

* California Institute of Technology, 105-24, Pasadena, California 91125, USA

† Jet Propulsion Laboratory, 138-307, Pasadena, California 91109, USA

‡ Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 5300 Bonn 1, FRG

Apparent faster-than-light motion has been detected with very long baseline interferometry (VLBI) in the central regions of ~12 extragalactic radio sources^{1,2}. These are typically strong, core-dominated objects and, so far the only weak-cored classical double radio source to show the effect was 3C179 (ref. 3). Unified source models⁴⁻⁶, which invoke relativistic motion at a small angle to the observer's line of sight, predict that large superluminal motion should be rarely seen in a randomly oriented sample of such objects. These models have motivated statistical studies of quasar samples⁷⁻⁹, designed to detect superluminal objects and, subsequently, to compare the distribution of velocities with model predictions. Here we report the detection of superluminal motion

in the quasar 3C263. The milliarc second structure of this source consists of two nearly unresolved components, whose separation is increasing at a rate of $0.06 \pm 0.02 \text{ mas yr}^{-1}$. This corresponds to an expansion speed of $2.7 \pm 0.9 c$ ($H_0 = 55 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q_0 = 0.05$). This quasar is the weakest superluminal source found so far, and there are indications that superluminal motion occurs frequently in this class of object.

The quasar 3C263 (=1,137+660; $m_b = 16$, variable¹⁰; $z = 0.652^{11}$) was earlier mapped with arc second resolution at 5 GHz¹¹, at 2.7 and 8.1 GHz¹², and at 1.42 GHz¹³. It has a steep overall spectrum¹⁴ and consists of a weak core (~130 mJy at 5 GHz) with inverted spectrum and two dominant steep-spectrum lobes with separation of ~45'' in position angle (p.a.) 110°. There is no evidence for radio variability of the total source flux density at 5 GHz¹⁵ or at 10.7 GHz¹⁶, nor is variability of the central component detected at 15 GHz (J. F. C. Wardle, personal communication). Earlier VLBI observations of the core at 5 GHz¹⁷ and at 10.7 GHz⁷ showed an extended core structure aligned with the outer double lobes.

We observed the core of 3C263 at three epochs with a transatlantic VLB array operating at a frequency of 10.7 GHz (left-hand circular polarization). These observations were made at intervals of ~1 yr, each time with up to five antennas (Table 1). The Mark III recording system¹⁸ was used in mode A or B (56 or 28 MHz bandwidth). Data were recorded for ~12 h in scans

Table 1 Epochs of observation and modelfit parameters

Epoch	Stations	Component	S (mJy)	R (marcs)	p.a. (deg)	Axis (marcs)*
1982.75	BKGO	W	95^{+20}_{-10}	0	0	<0.2
		E	74 ± 20	0.48 ± 0.02	110 ± 2	<0.3
1983.78	BKO	W	86^{+15}_{-5}	0	0	<0.2
		E	45^{+15}_{-5}	0.53 ± 0.02	108 ± 2	<0.3
1984.94	BLKGO	W	103^{+15}_{-5}	0	0	<0.1
		E	52 ± 10	$0.61^{+0.02}_{-0.04}$	109 ± 2	0.5 ± 0.1

B, Max-Planck-Institut für Radioastronomie, Effelsberg, FRG; L, Istituto di Radioastronomia, Medicina, Italy; K, Haystack Observatory, Westford, Massachusetts USA; G, National Radio Astronomy Observatory, Green Bank, West Virginia, USA; O, Owens Valley Radio Observatory, Big Pine, California, USA. Errors given were determined by visual inspection and reduced χ^2 values for the agreement between model and observed visibilities.

* Axis is FWHM of circular gaussian components. At epoch III, component E is an elliptical gaussian, with axial ratio <0.7 and p.a. $103 \pm 20^\circ$.

of either 13 min each $\frac{1}{2}$ h or 6.5 min each $\frac{1}{4}$ h, with gaps used for tape rewinding and calibration measurements. The data correlation was carried out mainly on the MPIfR Mark III processor in Bonn, West Germany, and partly on the Haystack Mark III processor in Westford, Massachusetts. The data were then fringe-fitted and coherently averaged to integration times of 2–6.5 min. The correlation coefficients were calibrated by using system temperatures, peak sensitivities, and gain curves for each telescope. The r.m.s. errors for 2-min integrations were typically ~5 mJy for amplitudes and <5° for phases.

Hybrid maps^{19,20} were made for all epochs in a standard fashion⁷ (Fig. 1), showing a two-component structure oriented in p.a. ~110°. These images are 30% 'superresolved' through the use of a circular restoring beam of FWHM of 0.35 marcs which is 30% smaller than the conventional beam²¹. However, this superresolution is permissible²¹ and the images shown (especially component separations) are consistent with results obtained independently from model-fitting and maximum-entropy mapping^{21,22}. Structure parameters, determined from fitting two gaussian components to the data, are listed in Table 1. The component separation R changes from 0.48 ± 0.02 marcs (epoch I) to 0.61 ± 0.03 marcs (epoch III). In these models, the weaker component (E) is oriented in p.a. $110^\circ \pm 2^\circ$. Although the absolute flux densities S of the two components carry large formal errors (20–30%), the change in relative flux density suggests a decrease of ~35% for the weaker component. It also appears that the weaker component is expanding and elongated along the source axis at the third epoch. The total flux density of the models lies close to an extrapolation to 10.7 GHz of the flux density of the core (~170 mJy), and therefore suggests that we are not missing significant structure.

The compact structure in 3C263 is typical for the class of cores in steep-spectrum, lobe-dominated quasars^{7,9}. There is an apparent 'core-jet' structure a few parsecs in extent, as is commonly found in flat-spectrum, core-dominated quasars^{23–25}. The 'core-jet' classification requires component spectra that have not yet been obtained, but the fading, expansion, and elongation of component E lead us to identify it tentatively as the 'jet'. The quasar 3C263 also shows an 'orientation memory' preserving the position angle (~110°) from the parsec to kiloparsec scale. Hence it is a distant, lobe-dominated quasar, which shows the good alignment first noted for nearby, lobe-dominated radio galaxies, but stands in contrast to the misalignments in distant, core-dominated quasars²⁶.

Our primary result is the detection of an increase in separation between the core (W) and jet (E) components. As measured between epochs I and III, the proper motion is $\mu = 0.06 \pm 0.02$ marcs yr⁻¹. We can derive an apparent relative transverse velocity using the relation $v_{\text{app}}/c = \mu s(1+z)$, where z is the redshift and s the scale of light years per marcs²⁷. This yields $v_{\text{app}} = 2.7 \pm 0.9c$ ($H_0 = 55 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q_0 = 0.05$), qualifying 3C263 as a superluminal source; in fact, it has the weakest core of all sources found so far to be superluminal². In the relativistic jet model of Scheuer & Readhead⁴, our results require a minimum bulk Lorentz factor $\gamma = 2.9 \pm 0.8$ (with a

corresponding angle to the line of sight of $\theta \sim 20^\circ$), which is compatible with the values of γ suggested by the statistics of core strength^{4,6} (the maximum permissible θ , in the limit $\gamma \rightarrow \infty$, is 41°). If we assume, however, that the overall source axis is oriented at $\theta = 20^\circ$, then the deprojected angular size is ~130". This places 3C263 above the upper envelope of the distribution of largest angular size against redshift^{28,29}. A value $\gamma \geq 5$ would allow $\theta \geq 37^\circ$, relieving the difficulty with the large overall source size (this difficulty could be avoided by choosing a larger value of H_0 instead). No conclusions are warranted on the issue of one-sided against two-sided jets, however; Doppler diminution of a hypothetical counterjet with $\gamma = 2.9$ would easily conceal it from view, given our limit on the jet:counterjet brightness ratio (≥ 15 –30, based on a 5σ upper limit for the counterjet).

It is worthwhile to contrast the properties of 3C263 with those of 3C179³. The former shows relatively slower motion, is of larger projected linear size and contains a weaker central component than the latter, which has the strongest core in the sample of Zensus & Porcas⁹. One would, to first order, expect such differences from orientation effects assuming that both sources are intrinsically similar (in overall size, core strength and Lorentz factor) and that 3C263 has a larger angle to the line of sight.

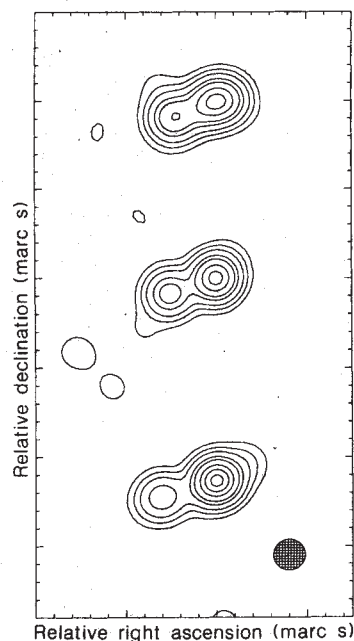


Fig. 1 Hybrid maps of the central component of 3C263 at epochs I (1982.75), II (1983.78) and III (1984.94), displayed from top to bottom. The vertical spacings correspond to the time intervals between the epochs; north is up, east is left; scale is 0.2 marcs per tick; contours are -2, 2, 5, 10, 20, 30, 50, 70, 90 mJy per beam. The circular restoring beam of FWHM = 0.35 marcs is shown cross hatched.

Contributing to the present inability to discriminate among various models is the uncertainty in H_0 . For randomly oriented sources in which each object has two intrinsically symmetric, but oppositely directed, relativistic jets, only one jet will be observed because of Doppler beaming. In a survey of such objects, the lower end of the velocity distribution should lie at $\geq c$. Therefore, such a survey could be used, in principle, to determine an upper bound to H_0 .

We note that the detected superluminal motion in 3C263 and 3C179, as well as preliminary results on more objects of this class^{7,9}, indicate that superluminal motion is a common phenomenon among cores of quasars with extended lobes. This trend will have to be accounted for in any successful model for superluminal radio sources.

We thank the staff at each observatory for their assistance, and Drs Marshall Cohen, Anthony Readhead and Arno Witzel for useful criticisms. The data from epoch I were the first to be correlated with the MPIfR Mark III processor and we give special thanks to all who contributed to its successful operation, especially Dr Alan Whitney and Herr Horst Blaschke. NRAO is operated by Associated Universities Inc., under contract with the NSF. This work was in part supported by the NSF through grants to the Owens Valley Radio Observatory. D.H.H. is an NRC-NASA Research Associate.

Received 13 August; accepted 16 October 1986.

1. Cohen, M. H. & Unwin, S. C. *IAU Symp.* **110**, 95–103 (1984).
2. Porcas, R. W. in *Active Galactic Nuclei* 20–49 (ed. Dyson, J. E.) (Manchester University Press, 1985).
3. Porcas, R. W. *Nature* **294**, 47–49 (1981).
4. Scheuer, P. A. G. & Readhead, A. C. S. *Nature* **277**, 182–185 (1979).
5. Blandford, R. D. & Königl, A. *Astrophys. J.* **232**, 34–48 (1979).
6. Orr, M. J. L. & Browne, I. W. A. *Mon. Not. R. astr. Soc.* **200**, 1067–1080 (1982).
7. Hough, D. H. Thesis, California Institute of Technology (1986).
8. Zensus, J. A. & Porcas, R. W. *IAU Symp.* **110**, 163–164 (1984).
9. Zensus, J. A. & Porcas, R. W. *IAU Symp.* **119**, 167–168 (1986).
10. Grandi, S. A. & Tift, W. G. *Publ. astr. Soc. Pacif.* **86**, 873–884 (1974).
11. Pooley, G. G. & Henbest, S. N. *Mon. Not. R. astr. Soc.* **169**, 477–526 (1974).
12. Owen, F. N., Porcas, R. W. & Neff, S. G. *Astr. J.* **83**, 1009–1020 (1978).
13. Shone, D. L. Thesis, Univ. Manchester (1985).
14. Kühr, H., Witzel, A., Pauliny-Toth, I. I. K. & Nauber, U. *Astr. Astrophys. Suppl. Ser.* **45**, 367–430 (1981).
15. Ekers, R. D., Fanti, R. & Miley, G. K. *Astr. Astrophys.* **120**, 297–301 (1983).
16. Seielstad, G. A., Pearson, T. J. & Readhead, A. C. S. *Publ. astr. Soc. Pacif.* **95**, 842–872 (1983).
17. Barthel, P. D. *et al. Astr. Astrophys.* **140**, 399–404 (1984); *Astr. Astrophys.* **151**, 131–136 (1985).
18. Rogers, A. E. E. *et al. Science* **219**, 51–54 (1983).
19. Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* **223**, 25–36 (1978).
20. Cornwell, T. J. & Wilkinson, P. N. *Mon. Not. R. astr. Soc.* **196**, 1067–1086 (1981).
21. Pearson, T. J. & Readhead, A. C. S. *A. Rev. Astr. Astrophys.* **22**, 97–130 (1984).
22. Gull, S. F. & Daniell, G. J. *Nature* **272**, 686–690 (1978).
23. Pearson, T. J. *et al. IAU Symp.* **119**, 163–164 (1986).
24. Eckart, A. *et al. Astr. Astrophys.* (in press).
25. Zensus, J. A., Porcas, R. W. & Pauliny-Toth, I. I. K. *Astr. Astrophys.* **133**, 27–30 (1984).
26. Readhead, A. C. S. *et al. Nature* **276**, 768–771 (1978).
27. Cohen, M. H. & Unwin, S. C. *IAU Symp.* **97**, 345–354 (1982).
28. Hooley, A., Longair, M. S. & Riley, J. M. *Mon. Not. R. astr. Soc.* **182**, 127–145 (1978).
29. Schilizzi, R. T. & deBruyn, A. G. *Nature* **303**, 26–31 (1983).

Activation of the C–C bond provides a molecular basis for structure sensitivity in metal catalysis

P. S. Kirlin & B. C. Gates

Center for Catalytic Science and Technology,
Department of Chemical Engineering,
University of Delaware, Newark, Delaware 19716, USA

The catalytic activation of hydrocarbon C–H and C–C bonds converts petroleum into fuels and chemicals. Understanding of the structural and electronic effects that control the catalysis has been a goal of researchers for decades. A focal issue concerns the number of metal atoms in a catalytic site required for C–C bond rupture^{1,2}. Metal surfaces catalyse alkane hydrogenolysis (C–C bond breaking), but the catalytic sites are poorly defined^{3,4}; soluble, structurally defined transition metal complexes are not known to catalyse this reaction. Here we present evidence demonstrating that isolated mononuclear rhenium complexes on the surface of MgO catalyse alkene hydrogenation, but not C–C bond rupture; in contrast, ensembles consisting of three of the same rhenium complexes catalyse both alkene hydrogenation and rupture of the C–C bond of cyclopropane. These results provide a molecular foundation for structure sensitivity in surface catalysis and point the way to design of multicentre catalytic sites on surfaces^{5–7}.

The problem of structure sensitivity in metal catalysis has been investigated with well-defined surfaces of single crystals of metal. Catalytic activities for alkane hydrogenolysis vary by two orders of magnitude from one crystal face of platinum to another, but all the faces have nearly the same activity for alkene hydrogenation^{2,8,9}. Evidence of structure sensitivity that is consistent with results of single-crystal studies has been obtained using supported metal catalysts of the kind used widely in industry, which consist of metal crystallites of various sizes and shapes on high-surface-area metal oxide supports. Rates of the structure-sensitive alkane hydrogenolysis (but not of the structure-insensitive alkene hydrogenation) depend on the average size of the metal crystallites^{3,4}. These results demonstrate that the activation of the C–C bond is sensitive to the microscopic

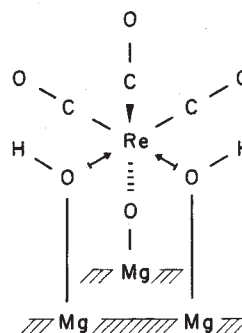


Fig. 1 Structure of the MgO-supported rhenium subcarbonyl, a catalyst precursor.

structure of the catalyst. However, exact structures of the catalytically active sites for hydrogenolysis are unknown and difficult to determine because of the lack of supported metal catalysts with unique particle sizes, the inseparability of individual active sites on single metal crystals, and the instability of highly unsaturated metal centres in molecular metal clusters, which might serve as models of metal surfaces.

Multiple dehydrogenation (C–H insertion) of cycloalkanes had been observed with soluble mononuclear rhenium complexes^{10–12}, and it is generally agreed that multiple dehydrogenation steps precede the scission of the C–C bond^{3,13} and that more than one metal atom is required in a catalytic site for C–C bond rupture^{3,4,8,9,13}. These results suggest that a well defined grouping of rhenium atoms (an ensemble), with one or more labile ligands per metal atom, may catalyse dehydrogenation of alkanes through C–H insertion, leading to catalysis of the rupture of the C–C bond on adjoining metal centres. We have prepared catalysts designed to incorporate ensembles consisting of three rhenium centres by reaction of the metal cluster $[H_3Re_3(CO)_{12}]$ with a high-surface-area porous support, MgO. For comparison, catalysts designed to incorporate single, isolated rhenium centres were made by reaction of $[HRe(CO)_5]$ with MgO. The interaction of the rhenium complexes with the basic MgO surface results in deprotonation (dissociative chemisorption) to form the adsorbed anions, $[H_2Re_3(CO)_{12}]^-$ or $[Re(CO)_5]^-$, respectively¹⁴. These anions are surface-bound intermediates; on heating, they are oxidized by surface –OH

Table 1 Activities of magnesia-supported rhenium catalysts

Catalyst precursor	Cyclopropane conversion							
	Propene hydrogenation		Isomerization to give propene		Reaction to give propane*		Hydrogenolysis to give methane and ethane	
	$10^3 \times \text{TOF}$ (s^{-1})†	E_a (kJ mol $^{-1}$)‡	$10^3 \times \text{TOF}$ (s^{-1})§	E_a (kJ mol $^{-1}$)‡	$10^3 \times \text{TOF}$ (s^{-1})	E_a (kJ mol $^{-1}$)‡	$10^3 \times \text{TOF}$ (s^{-1})	E_a (kJ mol $^{-1}$)‡
HRe(CO) $_5$ /MgO	35¶	57	#	#	#	#	#	#
H $_3$ Re $_3$ (CO) $_{12}$ /MgO	21¶	51	#	#	0.37¶	73	0.47¶	95
Re 0 /MgO	100**	31	0.18¶	15	870**	56	260**	56
Re 0 -CO/MgO††	0.063‡‡	70	0.12¶	73	0.12‡‡	124	0.18‡‡	143
Re metal powder	—	—	—	—	—	52 ¹⁵	—	52 ¹⁵

* Hydrogenolysis and/or slow isomerization followed by fast hydrogenation (see Fig. 2).

† Turnover frequencies (TOF) normalized to 353 K, 0.1 MPa, H $_2$ /C $_3$ = 8 (molar).

‡ Apparent activation energy $\pm$ 5 kJ mol $^{-1}$.

§ Turnover frequencies normalized to 493 K, 0.1 MPa, H $_2$ /C $_3$ = 1 (molar).

|| Turnover frequencies normalized to 473 K, 0.1 MPa, H $_2$ /C $_3$ = 9 (molar).

¶ Maximum rate.

Reaction undetectable.

** Initial rate, calculated on basis of assumption that all rhenium atoms exposed; value may be underestimated because conversion was greater than differential, too large to determine rate directly as conversion divided by inverse space velocity.

†† CO-poisoned metal on the support.

‡‡ Initial rate after CO poisoning.

groups of the MgO to form, respectively, groupings inferred to consist of three rhenium complexes (ensembles) or isolated rhenium complexes.

These surface-bound complexes have been characterized spectroscopically and inferred to have the composition [Re(CO) $_3$ {O-Mg}{HO-Mg} $_2$], where the curly brackets designate groups at the surface of the MgO (Fig. 1)¹⁴. The spectroscopic techniques used to characterize the structure include infrared, Raman, inelastic electron tunnelling, and ultraviolet-visible. The details of the surface chemistry will be published elsewhere with extended X-ray absorption fine structure (EXAFS) spectra that confirm the structure of Fig. 1 and support the inference that ensembles consisting of three rhenium complexes are formed on the surface from the trirhenium precursor. Here we compare the catalytic properties of isolated rhenium complexes, the ensembles, and the conventional form of an oxide-supported rhenium catalyst, namely, small, nonuniform crystallites of rhenium metal on the MgO surface.

The catalytic properties were investigated for a structure-insensitive reaction, the hydrogenation of propene, and a structure-sensitive conversion, the isomerization/hydrogenolysis of cyclopropane. This set of reactions, summarized in Fig. 2, has been used frequently to probe the nature of catalytic surfaces.

The activities of the several catalysts for propene hydrogenation (Table 1) confirm the known structure-insensitivity of the reaction. The close agreement between the apparent activation energies for propene hydrogenation on the catalysts derived from [HRe(CO) $_5$] and [H $_3$ Re $_3$ (CO) $_{12}$] (Table 1) is consistent with the observation that the rhenium complex shown in Fig. 1 was present on the surface of each catalyst, as determined by infrared spectra of the functioning catalysts. (These complexes are catalyst precursors; the precise structures of the catalytic intermediates are unknown.) The lack of activity of these catalysts for isomerization of cyclopropane to give propene suggests that groupings of more than three metal atoms are required for this reaction or, more likely, that zero-valent rhenium constitutes the catalytically active sites.

The data for the cyclopropane conversion (Table 1) demonstrate that an isolated rhenium complex, although it activates H $_2$ and catalyses alkene hydrogenation, does not catalyse rupture of the C-C bond; in contrast, an ensemble of three of these rhenium complexes is catalytically active for opening the cyclopropane ring to give propane, along with methane and ethane. As expected from the literature¹⁵, the larger groupings of metal

atoms of the rhenium metal crystallites also catalyse these reactions. On the basis of the data reported here, it is not possible to distinguish between the contributions of the two pathways to give propane (Fig. 2).

These results demonstrate a striking dependence of catalytic activity on the number of rhenium complexes constituting a catalytic site. Both the isolated [Re(CO) $_3$ {O-Mg}{HO-Mg} $_2$] and an ensemble of three of these complexes catalyse alkene hydrogenation. The ensembles, in contrast to the isolated complexes, also catalyse cyclopropane hydrogenolysis, and in this respect they resemble dispersed aggregates of rhenium partially poisoned with CO. Neighbouring rhenium centres are required for the hydrogenolysis. These results indicate how the use of organometallic precursors will allow the design of catalytic sites of varying numbers of metal atoms to allow precise control of selectivity in catalytic conversions.

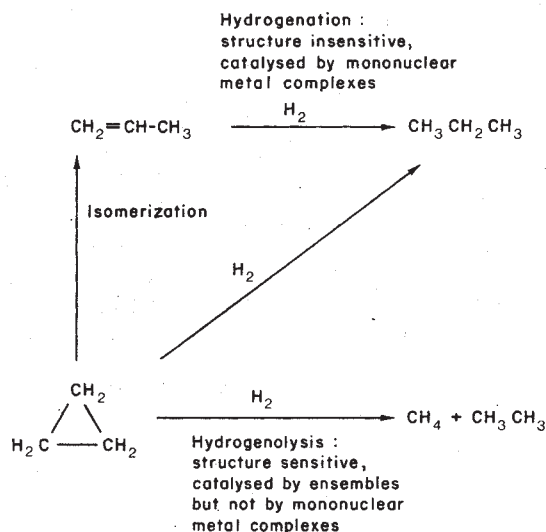


Fig. 2 Hydrocarbon conversion reactions used to probe the nature of the catalytic sites. Metals (including rhenium) or molecular ensembles on a support catalyse the hydrogenolysis of cyclopropane to give methane and ethane, but this reaction is not observed with the MgO-supported mononuclear rhenium complexes (see Table 1).

This research was supported by the Petroleum Research Fund, administered by the American Chemical Society, and by the Exxon Education Foundation.

Received 9 July; accepted 14 October 1986.

1. Boudart, M. *Adv. Catalysis* **20**, 153-166 (1969).
2. Somorjai, G. A. & Carrazza, J. *Ind. Engng Chem., Fundam.* **25**, 63-69 (1986).
3. Gault, F. G. *Adv. Catalysis* **30**, 1-95 (1981).
4. Van Broekhoven, E. H., Schoonhoven, J. W. F. M. & Ponec, V. *Surf. Sci.* **156**, 899-910 (1985).
5. Lisitsyn, A. S., Golouin, A. V., Kuznetsov, V. L. & Yermakov, Yu. I. *J. Catal.* **95**, 527-538 (1985).
6. Psaro, R. & Ugo, R. *J. molec. Catal.* **21**, 331-351 (1983).
7. Psaro, R. *et al. J. organometal. Chem.* **213**, 215-247 (1981).
8. Gillespie, W. D., Herz, R. K., Petersen, E. E. & Somorjai, G. A. *J. Catal.* **70**, 147-159 (1981).
9. Herz, R. K., Gillespie, W. D., Petersen, E. E. & Somorjai, G. A. *J. Catal.* **67**, 371-386 (1981).
10. Newman, L. J. & Bergman, R. G. *J. Am. chem. Soc.* **107**, 5314-5315 (1985).
11. Bandy, J. A., Cloke, G. N., Green, M. L. H., O'Hare, D. & Prout, K. *JCS Chem. Commun.* 240-242 (1984).
12. Baudry, D., Ephritikhine, M. & Felkin, H. *JCS Chem. Commun.* 606-607 (1982).
13. Rooney, J. H. *J. molec. Catal.* **31**, 147-159 (1985).
14. Kirtlin, P. S. *et al. J. phys. Chem.* **90**, 4882-4887 (1986).
15. Wallace, H. F. & Hayes, K. E. *J. Catal.* **18**, 77-82 (1970).

Acoustic emission during polymer crystallization

A. Galeski*[‡], L. Koenczoel^{†‡}, E. Piorkowska*[‡] & E. Baer§

* Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, 90-362 Lodz, Poland
[†] Fraunhofer Institute für Werkstoffmechanik, D-7800 Freiburg, FRG

§ Department of Macromolecular Science, Case Western Reserve University, Cleveland, Ohio 44106, USA

Ultrasonic emission was found to accompany several physical processes involving stress release, for example, microfracture in solids¹ or cavitation and bubble collapse in liquids². We here report the first observation of acoustic emission during crystallization. Acoustic emission was found to occur during the spherulitic crystallization of polymers from melt. The source of acoustic waves was identified as an abrupt stress release in regions of the melt occluded by spherulites during crystallization. The stress buildup in occluded areas is a result of density changes during crystallization. When the level of stress approaches the limit of the melt cohesion cavitation occurs, the stress is released and sound emitted. Because crystallization, in the form of aggregates growing simultaneously from many nuclei, always leads to the formation of occluded areas, and because the density of the melt or glass from which these aggregates grow is always lower than the crystallized matter, stress always arises. Depending on its level, the stress is either released by cavitation or remains frozen in the material. Thus acoustic emission is also expected to accompany crystallization in other, non-polymeric materials.

Crystallization of several polymers was investigated: isotactic polypropylene, poly(methylene oxide) and linear polyethylene. An acoustic emission analyser, Physics Acoustic Corporation apparatus PAC 3000/3004, with wide frequency band sensor (10 kHz-1 MHz) was used. Samples of polymers in the form of pellets containing 0.9 g of material were attached to the sensor, melted at temperatures above their melting points and then crystallized isothermally in a silicon oil bath at lower temperatures. The crystallization was conducted at several temperatures: 89-131 °C for isotactic polypropylene; 140-160 °C for poly(methylene oxide); 80-126 °C for high density polyethylene. During the entire crystallization process the acoustic emission was recorded. Non-crystallizing polymers such as atactic poly-

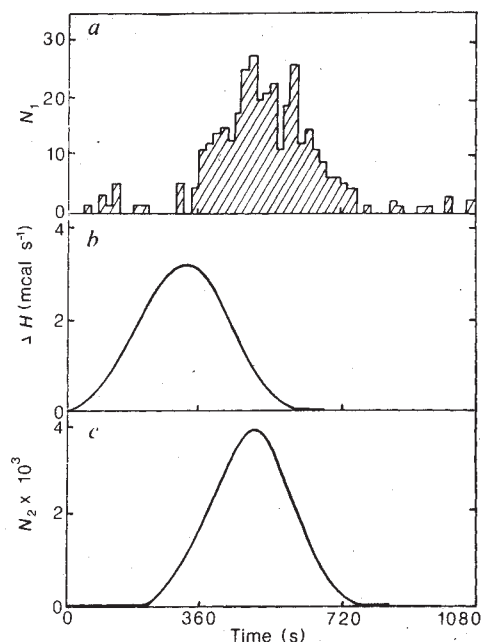


Fig. 1 *a*, Number of acoustic events plotted against time during isothermal crystallization of isotactic polypropylene at 120 °C. *b*, Heat of fusion generated during isothermal crystallization of isotactic polypropylene at 120 °C plotted against time. *c*, The time distribution of the formation of quadruple boundary points during crystallization of isotactic polypropylene at 120 °C. Instantaneous spherulitic nucleation assumed.

styrene and polybutadiene were subjected to the same treatment, for comparison.

It was found that a strong acoustic emission occurs during crystallization of isotactic polypropylene and poly(methylene oxide). Figure 1*a* is a typical diagram of the number of acoustic events plotted against time. The peak of arrivals of acoustic signals shifts towards longer time with increasing temperature of crystallization. The total number of recorded signals depends also on the temperature of crystallization. Few acoustic signals were observed for crystallization at low supercooling, more at higher supercooling and none or very few in quenched samples. An extensive modification of acoustic emission was obtained by the addition of a nucleating agent to the polymer (0.1 % talc). The non-crystallizing polymers do not show any acoustic emission if subjected to the same thermal treatment as the crystallizing polymers.

Comparison of the diagrams of acoustic events with data obtained from the isothermal crystallization of 10-mg samples of polymers in a differential scanning calorimeter (DSC-4) by Perkin-Elmer shows that most of the acoustic signals are generated close to the end of the crystallization process (see Fig. 1*b*, which plots the heat of fusion generated during crystallization against time).

Optical microscopy of thin sections of the samples reveals cavities inside the materials for which acoustic emission was detected during crystallization. The cavities are always located between spherulites, never inside spherulites. Linear polyethylene, which shows small spherulites and no cavities under the optical microscope produced very little acoustic emission during crystallization.

It becomes evident that ultrasonic emission is directly related to the crystallization and, more precisely, to the spherulite growth and the presence of cavities in the material. The investigated polymers crystallize in a form of polycrystalline aggregates, spherulites that grow radially from primary nuclei. During crystallization impingement of spherulites occurs, leading to the formation of interspherulitic boundaries. It was recently shown³ that the impingement of spherulites also causes another effect:

[‡] Present addresses: Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA (A.G., E.P.). Department of Macromolecular Science, Case Western Reserve University, Cleveland, Ohio 44106, USA (L.K.).

occlusion of regions of melt by spherulites. The crystallization of these regions proceeds without a supply of fresh melt to fill up the deficiency in the material caused by the density change during crystallization. Hence, the material can never be fully dense. It becomes obvious that the deficiency of the material leads to the buildup of stresses in the bulk of the sample. If the level of stress becomes higher than the limit of cohesion of the melt, cavitation occurs and the stress is abruptly released. It follows, that acoustic waves observed during crystallization, are brought about by fast stress release that accompanies cavitation in melt regions surrounded by growing spherulites.

Because the mathematical description of the formation of spherulitic structure is available⁴ one can correlate the acoustic emission with the formation of quadruple boundary points between spherulites where the cavities are found. On the basis of the mathematical description of spherulitic structure⁴ one can derive the time distribution of formation of quadruple boundary points in bulk samples for model modes of spherulitic nucleation. The time distribution for instantaneous nucleation expressed in half-time of crystallization is described by:

$$p(t)dt = 1.5 (\ln 2)^3 t^{8/2} e^{-t} dt \quad (1)$$

and plotted in Fig. 1c for isotactic polypropylene crystallized at 120 °C. This dependence reaches the maximum when the conversion degree of melt to spherulites is about 93%. From similar calculations for a spontaneous type of nucleation, it follows that the largest number of quadruple boundary points is formed when the degree of conversion is 95%. The time distribution of the formation of quadruple boundary points is in very good agreement with the time dependence of acoustic emission (compare Fig. 1a and Fig. 1c).

The change in acoustic emission after addition of the nucleating agent indicates the importance of spherulite number and size for the phenomenon. The increase in the number of spherulites increases the number of regions occluded during crystallization, but decreases their sizes. The cavity is formed within occluded region when the stress is large enough to overwhelm the cohesion of melt. The level of stress is determined by the mechanical properties of the crystallized matter, the compressibility of the melt, the difference in densities between the melt and the crystallized matter, and the size of the occluded area. If the occluded region is small, the stress does not reach the cohesion limit; it remains frozen in the bulk of the material and no acoustic emission is expected. Hence one can observe the formation of only a few holes and little acoustic emission for samples in which small spherulites are formed. This agrees with the observation of the drastic decrease in acoustic emission after addition of the nucleating agent and little acoustic emission during crystallization of polyethylene containing small spherulites and having low modulus. The changes in acoustic emission with the crystallization temperature reflect changes in cavitation due to variations in the number and sizes of the occluded regions.

The occlusion of melt regions by growing domains is a general feature of domain crystallization. Domain crystallization always leads to buildup of stresses in occluded regions and causes cavitation if stresses are high enough. Hence the cavitation between crystalline domains is a part of structure formation. The cavitation and then acoustic emission is also expected during crystallization of other, non-polymeric materials. The acoustic-emission technique may be used to monitor cavitation during domain-structure formation, which is important because cavities are weak spots in a material and decrease its mechanical performance.

Partially supported by the Fulbright-Hays program and Deutsche Forschungsgemeinschaft (Fulbright grant to A.G.; part support to L.K.)

Received 21 July; accepted 16 October 1986.

1. Acoustic Emission, Special Technical Publication 505, ASTM, Philadelphia, 1971; Grabec, I. & Peterlin, A. *J. Polymer Sci. Polym. Phys. Ed.* **14**, 651 (1976).

- Hammit, F. G. *Cavitation and Multiphase Flow Phenomena*, 1-6 (McGraw-Hill, New York, 1980).
- Galeski, A. & Piorkowska, E. *J. Polymer Sci. Polym. Phys. Ed.* **21**, 1299-1312, 1313-1322 (1983).
- Piorkowska E. & Galeski, A. *J. Phys. Chem.* **89**, 4700-4703 (1985); *J. Polymer Sci. Polym. Phys. Ed.* **23**, 1273-1748 (1985).

Fractal patterns from chemical dissolution

Gérard Daccord & Roland Lenormand

Dowell-Schlumberger, BP 90, 42003 Saint Etienne Cedex 1, France

The highly ramified patterns^{1,2} produced by the flow of a reactive fluid through a soluble porous medium have never been quantitatively described. The theoretical understanding of this phenomenon is limited to very simple conditions (such as the flow of a liquid through a capillary³) due to the complexity of the coupling between the chemical reaction and the fluid flow. We show here that the dissolution patterns (DP) obtained experimentally by injecting water through pure plaster are fractal, for different geometries of the samples. In two dimensions, these DP are remarkably similar to patterns associated with diffusion-limited aggregation⁴⁻⁶ (DLA), that is, dielectric breakdown⁷, viscous fingering^{8,9} and diffusion-limited polymerization¹⁰. In three dimensions, we compare them with DLA clusters grown in the same boundary conditions and find a good qualitative and quantitative similarity. These results should be of interest in different areas where chemical dissolution of porous media by a flowing fluid occurs, for example, in nature (the formation of caves) and in industry (in the oil industry where acids are routinely injected into oil reservoirs).

We used the model system plaster/water because its properties^{1,2} are well known and it is convenient for studying various geometries. Plaster samples were prepared by mixing pure water with reagent grade $\text{CaSO}_4 \cdot 0.5 \text{H}_2\text{O}$ in the ratio of 1/1.1 by weight. After setting in an appropriate mould, the plaster (60% porosity and $60 \mu\text{m}^2$ permeability) was saturated with water equilibrated with plaster and then pure water was injected at constant flow rate, the injection pressure P being recorded.

We used three types of geometry. Type I: two-dimensional radial consisting of a thin disk (typically 1 mm thick and 200-mm diameter). Water was injected at the centre. Type II: thick disks, made from cylindrical samples (50-mm diameter and 60-mm long) with a hole (2-mm diameter) drilled along their axes. In this case, water was injected radially from the central hole. Type III: cylindrical samples for which we varied the ratio of diameter/length (from 50 mm/30 mm to 25 mm/100 mm), the water being pumped linearly through one end face.

All the experiments produced highly ramified, tree-like structures, resembling patterns associated with DLA. We obtained mouldings by injecting a polymerizable resin or a low-melting point alloy (Wood's metal) into the DPs and subsequently dissolving the remaining plaster. Figure 1 displays typical DPs obtained for the three geometries studied. A remarkable feature of these structures is the decreasing width of the branches as we go from their origin (the injection point, line of surface) to their tip. Also characteristic is the effect of the flow rate: as it increases, the number of 'trees' increases together with the number of ramifications.

The two-dimensional DPs were easily quantified by digitizing the photographs and computing the density-density correlation functions⁴. We obtained a fractal dimension $d_f = 1.6 \pm 0.1$, in reasonable agreement with results from two-dimensional DLA. This fractal behaviour is observed to extend over at least three decades, in agreement with microscopic observations of the mouldings. Indeed, the minimum branch size of the mouldings is of the order of tens of micrometres and the overall dimension being typically 50-100 mm, we retrieve the three decades,

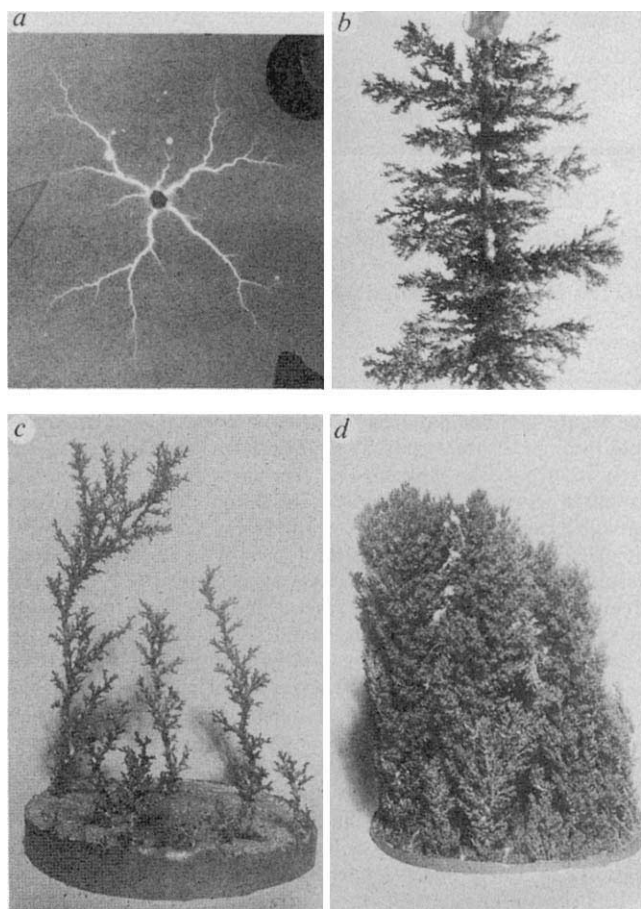


Fig. 1 Photograph of dissolution patterns obtained for different geometries. *a*, Two-dimensional pattern grown radially from a point observed by transparency. *b*, Three-dimensional pattern obtained after moulding a type III structure and dissolving the matrix. The pattern was grown from the inside cylinder which is 2-mm diameter and 50 mm long. *c*, *d*, Three-dimensional patterns obtained from type III experiments (outside radius 50 mm) at different flow rates, respectively $2.7 \text{ cm}^3 \text{ min}^{-1}$ and $40 \text{ cm}^3 \text{ min}^{-1}$.

This method of determining a fractal dimension is experimentally possible only for two-dimensional patterns. For three-dimensional DPs, we had to use an indirect method to get the fractal dimension, by analysing the injection-pressure curve. For all the radial experiments, we observed that the pressure P decreased linearly with $\ln t$. We interpreted this from Darcy's law in radial conditions: for a cylinder of homogeneous porous medium of outside radius R_0 , the pressure at a distance r from the axis is $P(r) \sim \ln(R_0/r)$, taking $P(R_0)$ as reference. Our experimental results are consistent with there being no pressure drop from $r=0$ to $r=R_c$; R_c being some characteristic dimension of the DP, and R_c varying with time according to a power law: $R_c \approx t^\alpha$. For type I experiments the inaccuracy of the measurement of the disk properties made the determination of α very inaccurate. We found $\alpha = 0.7 \pm 0.2$. For type II geometries, the accuracy was better: $\alpha = 0.65 \pm 0.07$ (we varied the flow rate over more than two decades, from 0.2 up to $40 \text{ cm}^3 \text{ min}^{-1}$, and found no significant difference in the fractal behaviour).

Brady and Ball¹¹ made comparable measurements during the electrodeposition of copper at constant voltage. They measured the current which, in their case, is proportional to the 'Smoluchowski radius' of the aggregate and assumed that it was proportional to the radius of gyration. In our case, making the same assumption (R_c proportional to the radius of gyration) would mean that for type I geometries $d_f = 1/\alpha$. Indeed, we find

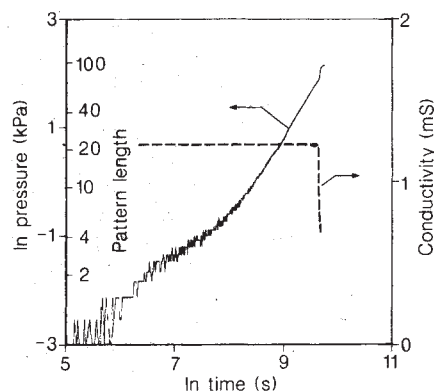


Fig. 2 Solid line, pressure curve for experiments with cylindrical samples (type III) displaying the crossover from three- to one-dimensional conditions. The second scale indicates the equivalent pattern length L_e , as a percentage of the sample length. Dashed curve, variation of the effluent water electrical conductivity.

a good agreement between the geometrical d_f (from the density-density correlation analysis) and the dynamic exponent $1/\alpha$. For type II experiments, our results have to be compared with diffusion limited aggregates grown on a fibre¹². In this case, our experimental value agrees very well with Meakin's exponent $\delta = 0.665$ of the root-mean square (r.m.s.) thickness of the deposit.

For type III experiments the pressure curve was less easily interpretable. Using the argument that there is no pressure drop in the sample until a characteristic distance L_c which varies with time according to $L_c \approx t^\alpha$, we should get a linear relation between $\ln(P - P_{\text{initial}})$ and $\ln t$. In fact, we observed two linear regions (Fig. 2): at short times, the slope is < 1 but when L_c reaches a value between 5 and 20% of the cylinder radius, it becomes $= 1$. We interpreted this as the consequence of the finite size of the sample: as long the structure size is small in comparison with the radius of the cylinder, the conditions can be approximated as three-dimensional. However, for longer injection times, when the pattern length is larger than this radius, the geometry can be considered as linear. Therefore, at the beginning we can compare our DP with diffusion limited deposits on a surface^{13,14} whose characteristic dimension is their r.m.s. thickness which scales with the total mass with the exponent ϵ : $1/\epsilon = D - d_f + 1$ (D is the Euclidean dimensionality). Then we would expect an initial slope of 0.57. Experimentally, the poor accuracy did not allow us to quantify exactly the initial behaviour (an error of 1% in P_{initial} changes the initial slope by 10–20%), but this prediction was at least qualitatively verified. On the other hand, for long times, we found a slope equal to unity in agreement with the linear conditions.

Similar behaviour was reported in the case of viscous fingering in a linear Hele Shaw cell (crossover from two- to one-dimensional) by Nittmann *et al.*¹⁵ who made static measurements of the effective fractal dimension for different cell widths. A few models of such finite size scaling exist (see refs 14, 16) but our data are not accurate enough for a quantitative comparison.

We also observed that the terminal part of the pressure curves displayed a discontinuity (Fig. 2). To understand this phenomenon, we put a conductivity cell at the outlet of the sample and measured the water electrical conductivity (which is a function of the concentration of ions): it displayed the same discontinuity as did the pressure, at the same time (Fig. 2*b*). This indicates that at this time t_c , the very tip of the DP reached the end face. This interface between the DP and the porous medium is also the reaction front, (where the water is fully saturated and no more reaction can occur) and is situated at a distance L_f from the source surface. At time t_c , we then have $L_f = L_0$, and the ratio pressure/initial pressure equals the ratio

L_c/L_f : we found values in the range from 0.6 and 0.9, indicating that L_c is significantly different from L_f . The determination of t_f for samples of different length gives access to the front velocity. We verified that for one-dimensional conditions, it is constant versus time and proportional to the flow rate.

We conclude that the chemical dissolution of plaster by water produces fractal patterns, and that these structures are closely related to DLA, for all the geometries studied. A more detailed discussion of the physical mechanism involved in this dissolution phenomenon and its relation with DLA will be published elsewhere¹⁷.

We thank Professor H. E. Stanley for many discussions and suggestions and J. Nittmann and Z. R. Lemanczyk for helpful criticisms.

Received 18 June; accepted 24 October 1986.

1. Ewers, R. O. Thesis, McMaster Univ., Hamilton, Ontario (1982).
2. Lee, P. B. & Mason, G. I. *Chem. E. Jubilee Symp. Ser.* 73, A13-21 (1982).
3. Levich, V. G. in *Physicochemical Hydrodynamics* (ed. Scriven, L. E.) 112-16 (Prentice-Hall, Englewood Cliffs, 1962).
4. Witten, T. A. & Sander, L. M. *Phys. Rev. Lett.* 47, 1499-1501 (1981).
5. Witten, T. A. & Sander, L. M. *Phys. Rev. Lett.* B27, 5685-5697 (1983).
6. Meakin, P. *Phys. Rev. A* 27, 604-607 (1983).
7. Niemeyer, L., Pietronero, L. & Weismann, H. J. *Phys. Rev. Lett.* 52, 1033-1036 (1984).
8. Måløy, K. J., Feder, J. & Jøssang, T. *Phys. Rev. Lett.* 55, 1885-1891 (1985).
9. Daccord, G., Nittmann, J. & Stanley, H. E. *Phys. Rev. Lett.* 56, 336-339 (1986).
10. Kaufman, J. H. *et al.* *Phys. Rev. Lett.* 56, 1932-1935 (1986).
11. Brady, R. M. & Ball, R. C. *Nature* 309, 225-229 (1984).
12. Meakin, P. *Phys. Rev. A* 27, 2616-2623 (1983).
13. Rácz, Z. & Vicsek, T. *Phys. Rev. Lett.* 51, 2382-2385 (1983).
14. Meakin, P. *Phys. Rev. B* 30, 4207-4214 (1984).
15. Nittmann, J., Daccord, G. & Stanley, H. E. *Nature* 314, 141-144 (1985).
16. Hentschel, H. G. E. & Deutsch, J. M. Preprint (Massachusetts Institute of Technology, 1986).
17. Daccord, G. *Phys. Rev. Lett.* (in the press).

Divergence of imaginary part of dielectric permittivity at the melting point in KCl

Jun Tateno

Japan Atomic Energy Research Institute, Tokai-mura, Naka-gun, Ibaraki-ken, Japan

The solid-liquid phase transition is a very common phenomenon, whose mechanism, although much studied, is not yet understood. A study of the properties of a solid having simple structure and simple bonding near the melting and the freezing point may be useful. We report the behaviour of the dielectric permittivity in KCl, a typical ionic crystal, in the microwave region in the vicinities of the melting and the freezing point. The imaginary part of the dielectric permittivity, ϵ'' , diverges at the melting and the freezing point; the real part, ϵ' , varies discontinuously. This supports a recent theory of melting, where the density fluctuations are emphasized as the precursors of melting or freezing.

The dielectric measurement in the microwave region is usually done by cavity method or by standing-wave method. In the latter method, which was introduced by Roberts & von Hippel¹, ϵ' and ϵ'' are determined by measuring the standing wave caused by the reflection in a sample. The sample was placed in a section of the waveguide, terminated by a reflecting plane, and the standing-wave ratio and the position of the minimum voltage were measured with a standing-wave detector. A single crystal of KCl, produced for optical purpose by Horiba Instrument Inc., was used as the sample. The sample cell (terminal part of the wave-guide) was made of platinum because of its high resistance to corrosion by KCl at high temperature. The sample cell was heated with an electric furnace, the temperature being controlled within the accuracy of ± 0.1 K. Measurements up to 800 °C were thus possible.

ϵ' and ϵ'' are given by¹

$$\epsilon' = (\lambda/\lambda_c)^2 - (\lambda/2\pi)^2(\alpha^2 - \beta^2) \quad (1)$$

$$\epsilon'' = 2\alpha\beta(\lambda/2\pi)^2 \quad (2)$$

$$\gamma = \alpha + i\beta \quad (3)$$

and

$$\frac{\tanh(\gamma d)}{\gamma d} = \left(\frac{\lambda_g}{2\pi d} \right) \left[\frac{\tan(2\pi y/\lambda_g) + ir}{ir \tan(2\pi y/\lambda_g) - 1} \right] \quad (4)$$

where r is the standing-wave ratio (the voltage ratio in node and antinode of the standing wave), y is the distance from the surface of dielectric to the first node, d is the thickness of the sample, λ and λ_g are the wavelength in free space and in air inside the waveguide, respectively, and λ_c is the cut-off wavelength. Using the measurement frequency of 9.21 GHz, we have $\lambda = 3.27$ cm, $\lambda_g = 4.75$ cm and $\lambda_c = 4.58$ cm. Because of vari-

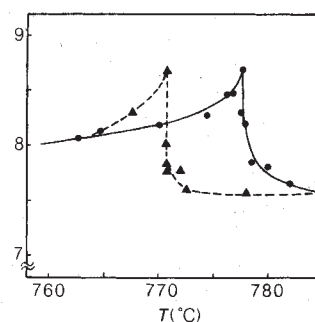


Fig. 1 Temperature dependence of ϵ' in KCl at 9.12 GHz: ●, heating; ▲, cooling.

ation of thickness of the sample with temperature, thermal expansion and volume change at the melting and the freezing point, and the meniscus of the liquid sample were ignored, for simplicity; the thickness was taken as constant ($d = 0.85$ cm).

Between room temperature and just below the melting point, ϵ'' of the KCl single crystal is nearly constant ($\epsilon'' = 0.007 \pm 0.001$ at 25 and 700 °C), whereas ϵ' increases with increasing temperature ($\epsilon' = 4.51 \pm 0.05$ at 25 °C and 7.82 ± 0.05 at 700 °C). These results in the single crystal are reproduced for the solid sample after freezing.

Temperature dependencies of ϵ' and ϵ'' of KCl in the vicinities of solid-liquid transition points are indicated in Figs 1 and 2. The value of ϵ' varies discontinuously at these points. The value of the imaginary part, ϵ'' diverges apparently at the transition points (Fig. 2). The disagreement between the temperatures for maximum values of ϵ'' in heating and cooling is due to the superheating and supercooling.

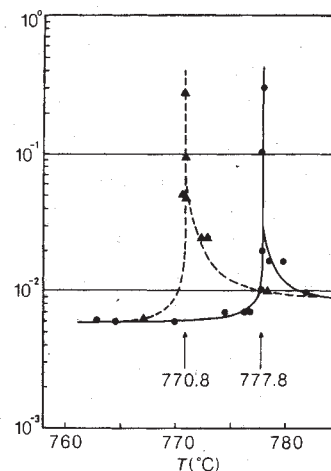


Fig. 2 Temperature dependence of ϵ'' in KCl at 9.12 GHz: ●, heating; ▲, cooling.

The present theories of melting can be classified into two groups. One, which is statistical and thermodynamical, is based on the principle that the free energies of both the phases are equal at the transition point. The other is based on elastic instability of the solid lattice, such as the shear modulus dropping to zero at the transition point². A particular mechanism causing the divergence of ϵ'' is not found in the above theories of melting. In the higher-order phase transitions, the susceptibility such as magnetic susceptibility or dielectric permittivity generally diverges at the transition point. This phenomenon is well explained by the theory of order-disorder phase transition. However, the present result is not explained by this treatment, because the solid-liquid phase transition is a first-order phase transition where the two phases coexist at the transition point.

Recently a new theory of melting has been proposed^{3,4,5}: the stability limit in solid phase is far above the melting point, so that the dynamical lattice instability is not related to melting⁵. This stipulates that the enhancement of density fluctuations in the solid lattice or in the liquid phase causes the solid-liquid phase transition. From this point of view, Schneider⁴ has shown that the freezing point or the stability limit of liquid is determined by $-1/\chi(Q) = kT/S(Q) = 0$, which is derived from the generalized minimum principle of free energy. Here, $\chi(Q)$ is the density-

response function and $S(Q)$ is the mean density fluctuation (structure factor) with the wavenumber Q . Therefore fluctuation is enhanced in liquid-solid phase transitions. On the other hand, statistical study of the fluctuations shows that the imaginary part of the dynamic susceptibility diverges in enhancement of fluctuations⁶. Some critical phenomena may be a typical example of this situation⁷.

In the microwave or the far-infrared region the dielectric nature is mainly related to the behaviour of ions (lattice polarization). Therefore, the present result of divergence in ϵ'' shows the enhancement of ion-density fluctuations in the solid-liquid phase transition. This also supports the new treatment of solid-liquid phase transition where density fluctuations are emphasized as the precursors of melting and freezing.

Received 28 May; accepted 24 October 1986.

1. Roberts, S. & von Hippel, A. *J. appl. Phys.* **17**, 610-616 (1964).
2. Boyer, L. L. *Phase Transitions* **5**, 1-48 (1985).
3. Schneide, T., Brout, R., Thomas, H. & Feder, J. *Phys. Rev. Lett.* **16**, 1423-1426 (1970).
4. Schneider, T. *Phys. Rev. A* **3**, 2145-2148 (1970).
5. Yaukalov, V. I. *Phys. Rev. B* **32**, 436-446 (1985).
6. Landau, L. D. & Lifshitz, E. M. *Statistical Physics* 2nd edn. 343 (translated by Sykes, J. B. & Kearsly, M. J.) (Pergamon, Oxford, 1968).
7. Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena* (Clarendon, Oxford, 1971).

Evidence from mantle xenoliths for an enriched lithospheric keel under the Outer Hebrides

Martin A. Menzies*, Alexander N. Halliday†, Zenon Palacz‡, Robert H. Hunter§, Brian G. J. Upton||, Pedre Aspen|| & Christopher J. Hawkesworth‡

* Department of Earth Sciences, Open University, Milton Keynes MK7 6AA, UK

† Scottish Universities Research Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

‡ Earth Sciences, Open University, Milton Keynes MK7 6AA, UK

§ Department of Earth Sciences, Cambridge University, Cambridge CB2 3EW, UK

¶ Grant Institute of Geology, Edinburgh University, West Mains Road, Edinburgh EH9 3JW, UK

|| Department of Geology, Royal Holloway and Bedford New College, University of London (RHBNC), Egham TW20 0EX, UK

Several independent lines of evidence suggest the presence of old regions in the Earth's upper mantle, particularly beneath stable cratons, which are enriched in incompatible elements such as barium and the light rare earth elements (LREE). Such 'lithospheric keels' are well documented beneath the Archaean of South Africa¹⁻⁴ and North America^{5,6}. Recent studies⁷ have shown the Lewisian of Scotland to be underlain by heterogeneous mantle, fragments of which are entrained by magmas that later intrude the crustal rocks of north-west Scotland. Here we present chemical and isotopic data from megacryst suites in these magmas, which point to mineralogical complexity in the lower crust and upper mantle. Moreover, variable trace-element ratios in the xenoliths are now recorded as extreme isotopic variations. The Sr and Pb isotopic compositions are similar to those of lamproites⁸ and micaceous kimberlites⁹ respectively. In contrast, the Nd isotopic ratios are highly variable ($\epsilon_{Nd} = -3.4$ to -32.6) and overlap with those observed in garnet inclusions⁹ in diamonds (South Africa) and lamproitic⁶-potassic⁸ volcanic rocks (USA). We suggest that the Hebridean lithospheric mantle has evolved in a manner totally unlike that of the mantle source regions tapped by Tertiary volcanic rocks erupted in the North Atlantic volcanic province.

The report⁷ of diverse suites of mafic and ultramafic xenoliths, the best evidence for heterogeneity in the Scottish mantle,

prompted investigation of the chemical constitution of the mantle below the Lewisian crust. We now report results for the Loch Roag monchiquite dyke, of Tertiary age (M. MacIntyre, personal communication), which intrudes Lewisian (Scourian) basement on Lewis in the Outer Hebrides. Inclusions in this dyke provide important constraints on the petrographic and geochemical nature of the Hebridean mantle. The major xenolith lithologies¹⁰ are mafic granulites, peridotites, pyroxenites and wehrlites, whilst the megacrysts comprise clinopyroxene, mica, magnetite, rutile, apatite, sanidine, anorthoclase, corundum and zircon.

The trace-element features of representative xenoliths are illustrated on a 'mantle normalized' diagram (Fig. 1) where the peridotite-pyroxenite data are compared with that of the host monchiquite at Loch Roag; lamproites from Smoky Butte Montana⁶, and micaceous kimberlites from the Finsch Mine, South Africa⁸. The ultramafic xenoliths are enriched in that they are characterized by significant concentrations of incompatible elements relative to compatible elements, in particular Ba, La, Ce and Nd. Moreover, the xenoliths have low Rb/Ba; low Rb/Sr and low K/Ti ratios, similar to the Smoky Butte lamproites⁶ and, in some respects, to the host monchiquite. In the case of several peridotites, it seems that certain trace-element concentrations are similar to the host magma while Rb, Ba, Nb, Sr, Zr, Ti define an enriched profile (dashed lines) at lower concentrations (Fig. 1). Possibly these peridotites comprise several components and have a complex origin.

The Sr and Nd isotopic compositions of the Loch Roag megacrysts, xenoliths and the host magma are presented in Table 1 and Fig. 2. The megacrysts (mica and apatite) and the peridotite-pyroxenite suite display a considerable range in $^{143}\text{Nd}/^{144}\text{Nd}$ (0.51247-0.51097) but a rather narrow range in $^{87}\text{Sr}/^{86}\text{Sr}$ (0.70415-0.70636). All the analyses except that for the apatite megacryst fall in the lower right-hand quadrant of the $\epsilon_{Nd}-\epsilon_{Sr}$ diagram (Fig. 2). The time-integrated response to the enrichment in LREE (low Sm/Nd) is a considerable variation in $^{143}\text{Nd}/^{144}\text{Nd}$ ratio, whilst the relatively low Rb/Sr ratio (0.017-0.128) of these xenoliths has generated a restricted range in $^{87}\text{Sr}/^{86}\text{Sr}$ ratio. These data are distributed in a vertical array on the $\epsilon_{Nd}-\epsilon_{Sr}$ diagram in a manner similar to that of the apatite-pyroxenite xenoliths found at Kiama, New South Wales¹¹, but very different from the bulk of spinel lherzolite xenoliths¹²⁻¹⁴ which tend to have Sr and Nd isotopic ratios similar to mid-ocean ridge basalts (MORB) and ocean island basalts (OIB). Isotopic data from the Loch Roag xenoliths

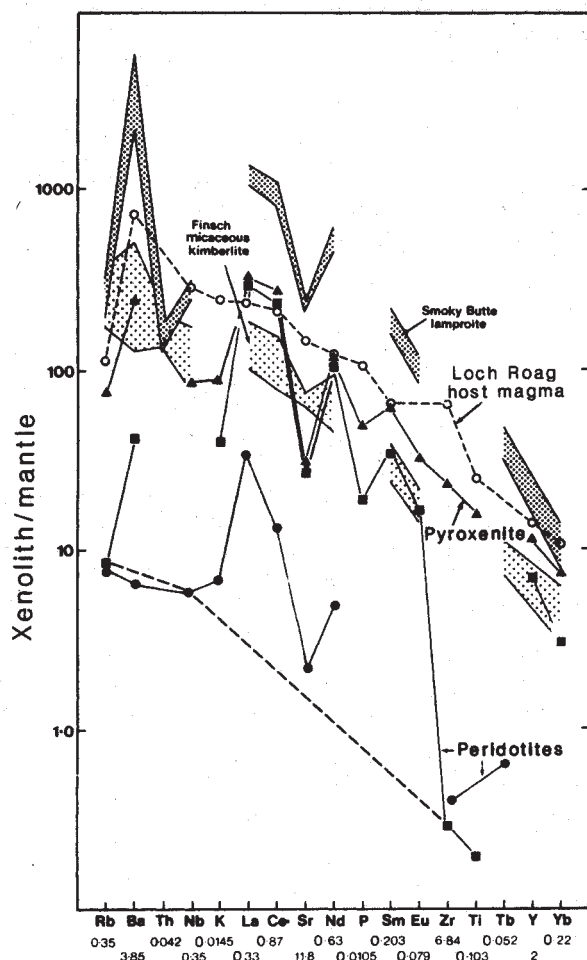


Fig. 1 Trace-element variation in Loch Roag xenoliths normalized to a mantle composition. Note that both the pyroxenite and the peridotites display a marked enrichment of incompatible elements relative to compatible elements. In particular, note the high concentrations of Ba and the LREE, and the depletions in K and Sr. These data are compared with the host monchiquite from Loch Roag, the Finsch micaceous kimberlite and the Smoky Butte lamproite. While the pyroxenite may be a derivative of such evolved melts, the peridotite may retain evidence of interaction with these melts in the form of a replacement metasomatic texture. The dotted profile may indicate the chemical characteristics of the peridotite before interaction with a melt. Trace-element data for the xenoliths and host magma from Menzies *et al.* (unpublished data). Note all normalization values (N. Rogers, personal communication) are for p.p.m. (parts per 10^6) except K, P, Ti which are per cent oxide.

overlap with Sr–Nd data from garnet inclusions found in megacrystic diamonds³ from South Africa (Fig. 2). These silicate inclusions record an ancient (~3,000 Myr) enrichment in LREE believed to have occurred deep in the lithosphere. The Loch Roag xenoliths have Sm–Nd model ages indicative of an enrichment in LREE some 1,000–1,500 Myr ago. Moreover, the mica megacrysts record Sm–Nd model ages of 1,800–2,000 Myr. The Loch Roag xenoliths with the most radiogenic Nd isotopic ratios (LR32; LR80; LR102) have isotopic characteristics comparable with garnet ilherzolites from South Africa^{1,2,4}. Volcanic rocks from the North Atlantic¹⁵ fall within the ‘mantle array’ located in the upper left-hand quadrant of an $\epsilon_{\text{Nd}}-\epsilon_{\text{Sr}}$ diagram (Fig. 2). Because these data are a reflection of the mantle composition below the ocean basins, it is apparent that the sub-Hebridean mantle as represented by the Loch Roag xenoliths, has evolved in response to different trace-element processes relative to the source of Atlantic MORB. However, there are continental volcanic rocks with Sr and Nd isotopic characteristics identical to

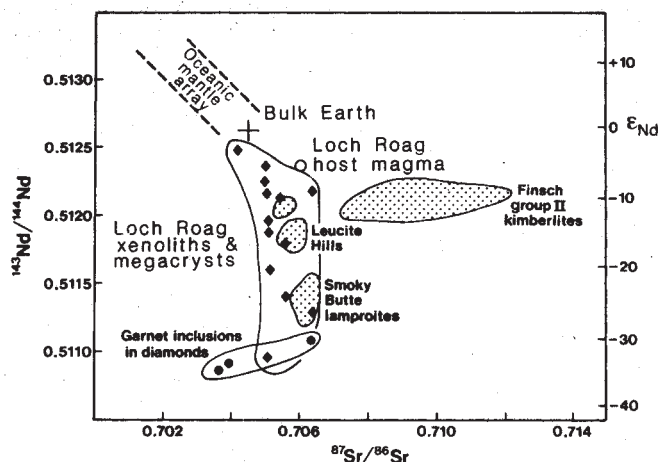


Fig. 2 $^{143}\text{Nd}/^{144}\text{Nd}$ plotted against $^{87}\text{Sr}/^{86}\text{Sr}$ for the Loch Roag xenoliths and megacrysts. The extremely narrow range in $^{87}\text{Sr}/^{86}\text{Sr}$ ratio and the heterogeneity in Nd isotopic composition reflect low Rb/Sr ratios and variable enrichments in the LREE. Shown for comparison are the data for garnet inclusions found in megacrystic diamonds⁹ in kimberlite pipes. The range in Nd and Sr isotopes compares well with that of the madupites and wyomingites (potassic volcanic rocks) of the Leucite Hills⁵ and the lamproites of Smoky Butte Montana⁶. The derivation of these potassic-ultrapotassic rocks may involve similar lithospheric mantle enriched in Ba, LREE and depleted in Rb/Sr. Oceanic data are taken from refs 26–28.

those of the Loch Roag peridotite–pyroxenite suite. These include the Smoky Butte lamproites from the western USA⁶, and the madupites and wyomingites from the Leucite Hills Wyoming⁵ (Fig. 2). The Loch Roag host magma is also shown in Fig. 2, as this has some of the trace-element and isotopic characteristics of the xenoliths. The distinctive characteristic of the Loch Roag xenolith suite is the vertical trend defined by the Nd and Sr data on an $\epsilon_{\text{Nd}}-\epsilon_{\text{Sr}}$ diagram. The similarity to diamond inclusions and to lamproites indicates that the origin of the lithospheric mantle below the Hebrides may have involved similar trace-element enrichment processes.

Pb isotopes (Table 2, fig. 3) were determined using techniques outlined elsewhere³³. The Loch Roag xenoliths show a narrow range in $^{206}\text{Pb}/^{204}\text{Pb}$ (17.0–17.36) and $^{207}\text{Pb}/^{204}\text{Pb}$ (15.44–15.50), with more variation in $^{208}\text{Pb}/^{204}\text{Pb}$ (36.9–37.57). The data overlap the field of micaceous kimberlites (Group II)⁸, but differ greatly from the Smoky Butte lamproites⁶ which otherwise share trace-element and Sr–Nd isotopic similarities. The Loch Roag xenolith data require evolution in an environment with low μ (and low Sm/Nd), like many Group II kimberlites^{6,8} and West Australian lamproites^{6,16}. No obvious correlation exists between Pb isotopes and Sr or Nd isotopes, implying that they may have evolved independently of each other. The Loch Roag xenoliths have distinct Pb isotopic compositions unlike oceanic rocks, but like those of lamproites and group II kimberlites, again demonstrating that evolution of these lithospheric fragments occurred in an environment different to that found beneath ocean basins and more like that occurring beneath the cratonic regions of North America, South Africa and Australia.

The ultramafic–mafic xenoliths and megacrysts entrained in the Loch Roag dyke represent fragments of chemically enriched lithosphere. The extreme heterogeneity in Nd isotopic ratios represents an integrated response, over some 1,000–1,500 Myr, to an enrichment in the LREE. The limited variation in $^{87}\text{Sr}/^{86}\text{Sr}$ ratio is presumably a reflection of the limited range in Rb/Sr measured in these xenoliths. In contrast, some of the mica megacrysts give older model Nd ages of 1,800–2,000 Myr, the significance of which is not fully understood. It is of interest, however, in that they could potentially point to ancient metasomatic events in the lithosphere. As lower crustal xenoliths are

Table 1 Rb, Sr, $^{87}\text{Sr}/^{86}\text{Sr}$, Sm, Nd, $^{143}\text{Nd}/^{144}\text{Nd}$ data for ultramafic/mafic xenoliths and megacrysts from Loch Roag, Lewis, Scotland

Sample	Description	Rb (p.p.m.)	Sr (p.p.m.)	$^{87}\text{Sr}/^{86}\text{Sr}$	Nd (p.p.m.)	Sm (p.p.m.)	$^{143}\text{Nd}/^{144}\text{Nd}$
LR21	Mica megacryst	211.7	385.0	0.70495 ± 4	2.87	0.33	0.51097 ± 4
LR22	Mica megacryst	236.8	363.3	0.70636 ± 9	1.83	0.27	0.51217 ± 3
LR23	Mica megacryst	255.5	291.4	0.70632 ± 4	1.72	0.18	0.51128 ± 4
LR32	Apatite megacryst	0.573	6636	0.70415 ± 4	738	106	0.51247 ± 2
LR64	Pyroxenite	9.83	353.3	0.70535 ± 4	31.0	5.71	0.51212 ± 2
LR80	Spinel lherzolite	3.01	23.5	0.70494 ± 10	1.83	0.32	0.51236 ± 2
LR81	Spinel lherzolite	2.62	154.8	0.70497 ± 3	67.5	7.16	0.51215 ± 2
LR95	Pyroxenite	25.87	370.8	0.70558 ± 3	(75.2)	12.5	0.51138 ± 2
LR101	Pyroxenite	2.486	80.44	0.70514 ± 3	23.1	3.94	0.51160 ± 2
LR102	Pyroxenite	—	231.5	0.70488 ± 3	(61.1)	13.20	0.51225 ± 3
LR(H)	Monchiquite host	40	1755	0.70597 ± 3	(72.8)	12.7	0.51231 ± 2

Most Rb/Sr obtained at SURRC using standard techniques. All Sm/Nd were obtained at the Open University except LR21–23 which used a modified version of established techniques³³ with Nd blanks <40 pg.

not mica-bearing, we believe that the micas represent disrupted mantle pegmatites. Basic granulite xenoliths¹⁷ at Loch Roag record Sm–Nd model ages at 2,500–3,500 Myr similar to those of the exposed Lewisian gneisses. The difference in the age between the felsic and ultramafic xenoliths may reflect a complex evolution for this region. Interleaving⁷ of ultramafic (possibly pyroxenitic) lower crust and felsic upper mantle due to magmatic underplating may explain the diverse xenolith population and the age differences.

Several features of the peridotite xenoliths indicate that they are perhaps multicomponent. Replacement-reaction textures due to metasomatism are visible in many lherzolites, a process that may explain the unusual aspects of the normalized trace-element diagrams. If 'mixing' between a peridotite and a melt or fluid has occurred, any age information is suspect. It is difficult to know to what extent we can interpret the isotopic characteristics of the xenoliths as inherently those of the lithosphere and not a mix of peridotite (lithosphere?) and a melt (asthenosphere?). Perhaps the vertical array (Fig. 2) results from mixing of Lewisian lithospheric peridotite ($^{143}\text{Nd}/^{144}\text{Nd}$ of 0.5108–0.5109 and $^{87}\text{Sr}/^{86}\text{Sr}$ of 0.703–0.704) and Tertiary asthenospheric melts ($^{143}\text{Nd}/^{144}\text{Nd}$ of 0.5120–0.5124, $^{87}\text{Sr}/^{86}\text{Sr}$ of 0.705–0.706). In other words, the xenoliths may have been modified by the processes that transport them to the surface. The xenolith data are then confined by two hyperbolae constructed between these endmembers at high Sr/Nd and low Sr/Nd ratios. Although this may explain some of the data it cannot explain the existence of pyroxenites (LR95) with low $^{143}\text{Nd}/^{144}\text{Nd}$ ratios and high concentrations of Nd. Furthermore, the isotope-trace-element correlations anticipated as a result of such a model do not exist in all cases. This aspect of the data warrants further investigation.

The available Pb isotopic data for the Loch Roag xenoliths points to a two-stage evolutionary history not unlike that invoked for the source region of micaceous kimberlites and lamproites^{6,8,16}. The mantle represented by the xenoliths has not been involved in MORB or OIB genesis, having a higher $^{207}\text{Pb}/^{204}\text{Pb}$ ratio and a lower $^{206}\text{Pb}/^{204}\text{Pb}$ ratio than MORB. Evolution of Hebridean mantle began as a result of a high U/Pb ratio (high μ); and at ~1,500 Myr (Sm–Nd model age) a fractionation event caused a decrease in U/Pb ratio (low μ) producing the present-day variation. Alternatively, the Pb data represent a binary mixture between an infinite choice of endmembers. If the lithospheric mantle beneath north-west Scotland had evolved in a manner similar to lithospheric mantle beneath the American and African cratons (as exemplified by the lamproite–kimberlite data (Figs 2 and 3)) subsequent mixing or hybridization of mantle with a melt similar in composition to the Smoky Butte lamproite, would account for the present position of the Loch Roag data on Pb/Pb diagrams (Fig. 3). In addition, this would explain the remarkable similarity between the Loch Roag mantle and the Smoky Butte lamproites. If this scenario is feasible, the actual 'age' of the mantle below Loch Roag is

difficult to establish. Sm–Nd model ages may indicate the time of mixing lithospheric mantle and lamproite melts, whereas the intercept age on the Pb growth curve of 1,000 Myr would be fortuitous. One can state that the Lewisian is underlain, albeit locally, by chemically and isotopically heterogeneous lithospheric mantle characterized by low $^{143}\text{Nd}/^{144}\text{Nd}$ (low Sm/Nd); low $^{87}\text{Sr}/^{86}\text{Sr}$ (low Rb/Sr) and low $^{206}\text{Pb}/^{204}\text{Pb}$ ratios. Note that several independent lines of evidence have suggested the existence and possible participation of enriched mantle source

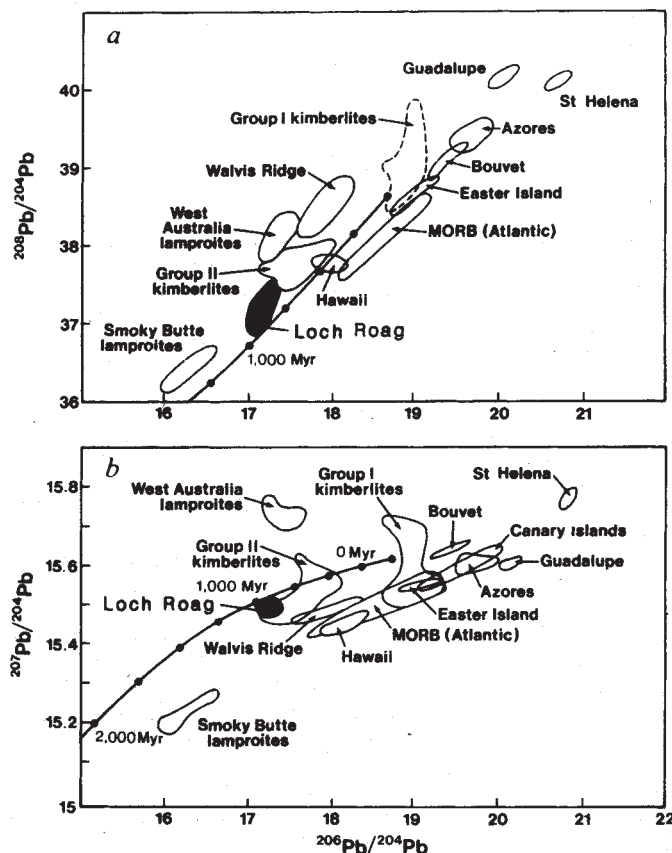


Fig. 3 *a*, $^{208}\text{Pb}/^{204}\text{Pb}$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$ and *b*, $^{207}\text{Pb}/^{204}\text{Pb}$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$ for the Loch Roag xenoliths. These data define arrays that intercept the Pb ore curve at 1,000–1,500 Myr. The similarity to Smoky Butte lamproites⁶ observed for Sr and Nd isotopes is not substantiated by the Pb data. The Loch Roag lithosphere has evolved more like that of the lithosphere believed to have been involved in the genesis of Group II kimberlites⁶. MORB, OIB and kimberlite data are taken from refs 8, 27, 29–31 and the Pb-ore growth curve is from ref. 32.

Table 2 $^{206}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ data for a selection of Loch Roag xenoliths and an apatite megacryst

Sample	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
LR32 (apatite)	17.359	15.497	37.412
LR64	17.139	15.457	37.151
LR80	17.209	15.476	36.911
LR81	17.259	15.464	37.059
LR95	17.301	15.490	37.250
LR101	17.231	15.453	37.343
LR102	17.348	15.440	37.570
LR(H)	17.096	15.486	36.962

Techniques are described elsewhere. Mean of 28 analyses of NBS 981 $206/204 = 16.886 \pm 10$; $207/204 = 15.427 \pm 13$; $208/204 = 36.494 \pm 35$ (2σ). Fractionation = 1.6‰ per AMU; blanks 0.5 ng. Errors on individual analyses are comparable to the errors on standard analyses ($\sim 0.1\%$).

regions during basaltic volcanism in the Siluro-Devonian and Tertiary period.

Thirlwall^{18,19} noted that relatively undifferentiated Ni-rich Siluro-Devonian volcanic rocks had high concentrations of incompatible elements in particular Sr, K, Ba, P and total LREE. More significant, however, are the marked geographical variations in the chemistry and isotopic compositions of volcanic rocks erupted through the Scottish crust. In particular, towards the north-west $^{143}\text{Nd}/^{144}\text{Nd}$ ratios of the lavas decrease. It has been suggested that mantle with a time-integrated LREE enrichment and radiogenic Sr isotopic compositions existed beneath the northwestern Highlands of Scotland, and that the lavas were derived from such an enriched source. How this unusual mantle was produced was unclear, but mixing of source regions was postulated. Interestingly, the increase in Ba, Sr, LREE, P and Nb towards the north-west coincides with the presence of old enriched mantle at Loch Roag, and mantle enriched in ^{87}Sr at Streap Comlaigh²⁰, as indicated by studies of mantle xenoliths. Earlier studies of the British Tertiary volcanic province (BTVP) alluded to the presence of spinel lherzolite mantle below north-east Skye which may have variable concentrations of incompatible elements like Rb, Sr, Ba, Nb, Th and the LREE. Extraction (and migration) of a lamprophyre²¹ (lamproite?) was believed to be the cause of such chemical heterogeneity. Mineralogical heterogeneity is also implied in the source regions of several Permian lamprophyres²². These are chemically equivalent to the nephelinites from Oahu, Hawaii and are believed to require retention of P (apatite), Nb (Nb rutile) and Ta (rutile?) in residual mantle phases. Such minerals are remarkably similar to the megacryst assemblages at Loch Roag (such as zircon $\pm$ Nb rutile $\pm$ apatite $\pm$ mica $\pm$ zircon). Beckinsale *et al.*²³ also believed that the early Tertiary of Skye needed a source contribution from heterogeneous mantle perhaps formed several billion years before the eruption of the BTVP. Although Moorbath and Thompson²⁴ agreed that the mantle below the Scottish Highlands may indeed be heterogeneous on various scales they stressed that the survival of any chemical and isotopic evidence of such heterogeneous sources depended on the extent of subsequent contamination²⁵ and fractionation of the volcanic rock in question.

It must be stressed that the evidence for mineralogical and chemical heterogeneity in the lithospheric mantle needs only time to produce isotopic variability. The Loch Roag xenoliths and xenoliths from Streap Comlaigh, Kilchatten and the Midland Valley²⁰ provide convincing evidence for a range of different types of intense metasomatic or enrichment processes within the mantle which have resulted in extreme isotopic heterogeneity. As a consequence, it cannot be assumed that the mantle beneath the north-west Highlands of Scotland has a Pb, Sr and Nd isotopic composition identical to MORB. Much of the underlying asthenosphere may have the isotopic characteristics of MORB and/or OIB but melting of heterogeneous lithospheric mantle and subsequent mixing with asthenospheric melts

could produce a series of volcanic rocks with elevated incompatible element concentrations and isotopic ratios not unlike those of Lewisian crust.

We thank the NERC for support through grants to the Open University and SURRC. The Isotope Geology Unit at SURRC is also supported by Scottish Universities. A.N.H. thanks P. J. Hamilton, J. Hutchison, F. MacLaren and J. Jocelyn for technical assistance. We thank Paul Taylor, Bob Thompson and Bill White for reviews, Matthew Thirlwall for comments and Sue Clay for assistance.

Received 22 May; accepted 28 October 1986.

1. Erlank, A. J. *et al.* in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) 221–311 (Academic, London, 1987).
2. Richardson, S. H., Erlank, A. J. & Hart, S. R. *Earth planet. Sci. Lett.* **75**, 116–128 (1985).
3. Hawkesworth, C. J., Erlank, A. J., Marsh, J. S., Menzies, M. A. & van Calsteren, P. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C. J. & Norry, M. J.) 111–138 (Shiva, Orpington, 1983).
4. Menzies, M. A. & Murthy, V. R. *Nature* **283**, 634–636 (1980).
5. Vollmer, R., Ogden, P., Schilling, J. C., Kingley, R. H. & Waggoner, D. G. *Contr. Miner. Petrol.* **87**, 359–368 (1984).
6. Fraser, K. J., Hawkesworth, C. J., Erlank, A. J. & Mitchell, R. H. *Earth planet. Sci. Lett.* **76**, 57–70 (1985).
7. Upton, B. J. G., Aspen, P. & Chapman, N. A. *J. geol. Soc. Lond.* **140**, 105–122 (1983).
8. Smith, C. B. *Nature* **304**, 51–54 (1983).
9. Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* **310**, 198–202 (1984).
10. Hunter, R. H. & Upton, B. G. J. in *Mantle Xenoliths* (ed. Nixon, P.) (Wiley, New York, in the press).
11. Menzies, M. A. & Wäss, S. Y. *Earth planet. Sci. Lett.* **65**, 287–302 (1983).
12. Menzies, M. A., Rogers, N., Tindle, A. & Hawkesworth, C. J. in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) 313–361 (Academic, London, 1987).
13. Zindler, A. & Jagoutz, E. *Geochim. cosmochim. Acta* (in the press).
14. Stosch, H., Carlson, R. W. & Lugmair, G. W. *Earth planet. Sci. Lett.* **47**, 263–271 (1980).
15. Carter, S. R., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Nature* **281**, 28–30 (1979).
16. McCulloch, M. T., Jaques, A. L. & Nelson, D. *Nature* **302**, 400–403 (1983).
17. Halliday, A. N. *et al.* *Scot. J. Geol.* **21**, 465–491 (1985).
18. Thirlwall, M. F. *J. geol. Soc. Lond.* **138**, 123–138 (1981).
19. Thirlwall, M. F. *Earth planet. Sci. Lett.* **48**, 27–50 (1982).
20. Menzies, M. A. & Halliday, A. N. *27th int. Geol. Cong. Abstr.* **5**, 348 (1984).
21. Thompson, R. N., Gibson, I., Marriner, G., Matthey, D. P. & Morrison, M. A. *J. Petrol.* **21**, 265–293 (1980).
22. Thompson, R. N. *Scot. J. Geol.* **18**, 49–107 (1982).
23. Beckinsale, R. D., Pankhurst, R. J., Skelhorn, R. R. & Walsh, J. N. *Contr. Miner. Petrol.* **66**, 415–429 (1978).
24. Moorbath, S. & Thompson, R. N. *J. Petrol.* **21**, 295–321 (1980).
25. Dickinson, A. P. *J. Petrol.* **22**, 155–189 (1981).
26. O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **34**, 13–22 (1977).
27. Cohen, R. S. & O'Nions, R. K. *J. Petrol.* **23**, 299–324 (1982).
28. Richard, P., Shimizu, N. & Allegre, C. *Earth planet. Sci. Lett.* **31**, 269–278 (1976).
29. Stille, P., Unruh, D. M. & Tatsumoto, M. *Nature* **304**, 25–29 (1983).
30. Richardson, S. H., Erlank, A. J., Duncan, A. R. & Reid, D. L. *Earth planet. Sci. Lett.* **59**, 327–342 (1982).
31. Sun, S. S. *Phil. Trans. R. Soc. A297*, 409–445 (1980).
32. Stacey, J. S. & Kramers, J. D. *Earth planet. Sci. Lett.* **26**, 207–221 (1975).
33. Palocz, Z. & Saunders, A. *Earth planet. Sci. Lett.* (in the press).
34. Menzies, M. A., Halliday, A. N., Hunter, R. H., MacIntyre, R. M. & Upton, B. G. J. *Australian Journal of Earth Sciences Special Publication, IV International Kimbelite Conference* (in the press).

Link between Archaean continent formation and anomalous sub-continental mantle

J. D. Kramers

Department of Geology, University of Zimbabwe,
PO Box MP 167, Mount Pleasant, Harare, Zimbabwe

The petrology and geochemistry of the subcontinental mantle from 70 to 160 km depth under Archaean cratons, as sampled by kimberlites, differs from the mantle sources of mid-ocean-ridge basalts (MORB) and hotspot volcanics. Here I propose that both the formation of this anomalous mantle and the origin of large masses of tonalitic-granodioritic crust in the early Archaean can be explained by a single open-system fractional crystallization model. If an extensive, partially molten layer existed at shallow levels in the pre-Archaean due to the production of large amounts of MORB-type magma above sites of mantle upwelling, the magma at loci of fractional crystallization would reach a steady state in major- and trace-element compositions which would be similar to that of Archaean granodiorites and tonalites. While this fractional crystallization process operated, the cumulate would chemically

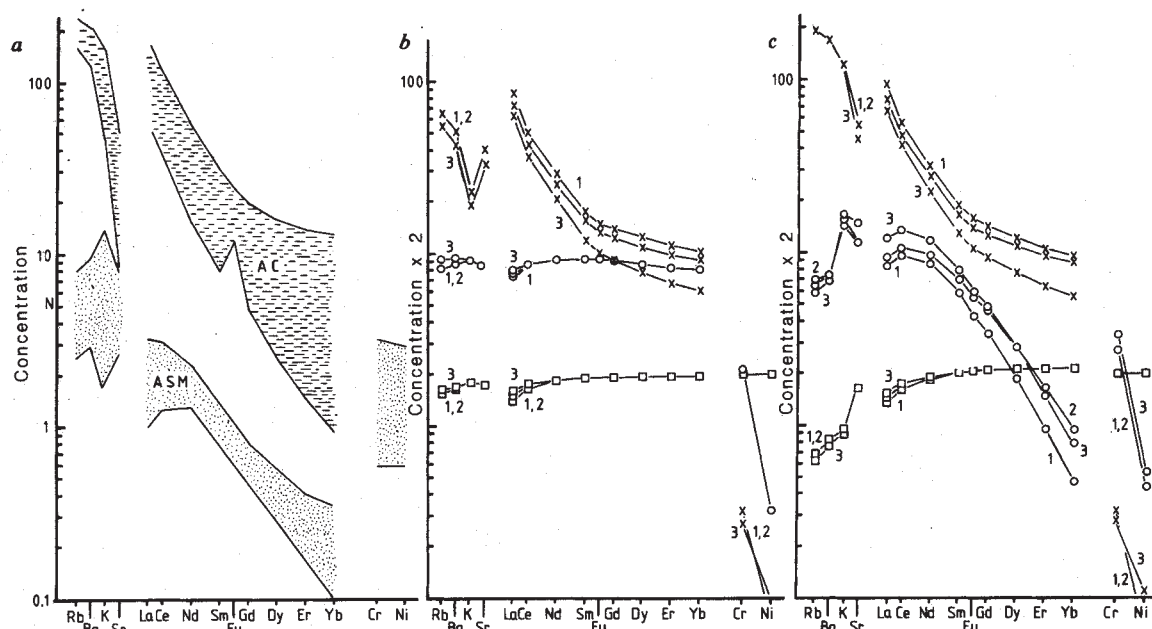


Fig. 1 Trace-element concentrations produced by the MARCY model (b, c) compared with observed values (a) for average continental crust^{17,18,28} (AC) and anomalous subcontinental mantle^{1,6,29-31} (ASM). In a, REE are normalized to C1 chondrite values; LIL elements, Cr and Ni are normalized to pre-crust mantle values²⁸ but doubled to retain coherence with C1-normalized REE patterns. In b and c, all concentrations are normalized against pre-crust mantle and doubled. For percentage fractionations, see Table 2. b, Steady-state trace-element patterns generated by the MARCY model. X, Fractionating melt body (protocrust); O, cumulate; □, mantle. All melting and crystallization occurs in batch (equilibrium) processes. MORB continuously admixed to fractionating body is effectively a product of 20% partial melting with an olivine + orthopyroxene residue. Fractionating body is 0.6% of total system, corresponding to upper mantle down to 600 km only participating in the process. Amphibole breakdown is not taken into account. c, Final trace-element patterns generated by MARCY model with continuous amphibole breakdown at the base of the cumulate, followed by episodic amphibole breakdown and garnet settling due to tectonic crustal shortening, modelled in 20 stages. At each stage, 5% of the existing amphibole breaks down, giving rise to (for 1 amphibole) 0.45 garnet + 0.45 clinopyroxene + 0.1 melt, the melt equilibrates with the total cumulate, the garnet is mixed into the mantle, the melt equilibrates again with the remaining cumulate and is then mixed into the crust. Note that LIL patterns now fit with a and subcontinental mantle patterns are approaching the right shape although abundances are too high. The apparent imbalance in HREE (mantle values >2.0) is caused by some low-REE residue of MORB production not being mixed back (not shown).

resemble a hydrated tholeiite. Pressure-induced breakdown of amphibole could cause partial re-melting of this cumulate, and this could lead to depletion resulting in a sub-continental mantle keel with chemical characteristics similar to those observed.

A large proportion of the mantle sampled by kimberlites appears to be made up of Mg-rich "cold coarse peridotites"¹ characterized by low temperatures of equilibration (850–1,100 °C)¹⁻³. These rocks are generally depleted in Fe, Ca, Al and other major elements more concentrated in basaltic melts than in their source rock^{1,4,5} and could thus be regarded as residues or cumulates. However, they often show enrichment in light rare-earth elements (LREE) (Fig. 1a), positive ϵ_{Sr} and negative ϵ_{Nd} values^{1,4-7}. Similarly anomalous ϵ_{Sr} and ϵ_{Nd} values for some kimberlites and continental alkali basalts have been attributed to equilibration with (or even origin from) such anomalous mantle material^{8,9}; although positive ϵ_{Sr} and negative ϵ_{Nd} values have been found in some oceanic volcanics, they are far less common and less extreme than those in continental volcanics⁹. This paradox between major element and incompatible element characteristics of the subcontinental lithosphere has generally been explained by invoking a metasomatic process following an earlier depletion^{5,7,10}. While metasomatic assemblages with hydrous minerals such as phlogopite and various amphiboles have been found, their trace-element abundance patterns differ from those of typical anhydrous anomalous sub-continental mantle, with much higher large-ion lithophile (LIL) element concentrations, lower Ce/Yb and higher La/Ce ratios^{10,11}. Such obvious (or "patent"¹⁰) types of metasomatism are young and can be related to volcanic events^{10,11}. For anomalous subcontinental mantle without patent metasomatism, evidence on ages is mainly circumstantial as it relies on Nd (refs 7, 12, 13) and Pb (ref. 14) model ages on peridotites^{7,14},

Table 1 Crystal-liquid partition coefficients^{17,32-34} used in the model calculations

	Olivine	Orthopyroxene	Clinopyroxene	Amphibole	Garnet
Rb	0	0.001	0.001	0.3	0.001
Ba	0	0.001	0.001	0.4	0.002
K	0	0.001	0.002	1.0	0.001
Sr	0	0.01	0.07	0.5	0.001
La	0	0.0015	0.15	0.2	0.01
Ce	0	0.002	0.2	0.4	0.03
Nd	0	0.005	0.35	0.7	0.2
Sm	0	0.01	0.5	1.1	0.8
Eu	0	0.016	0.55	1.2	1.3
Gd	0.001	0.025	0.6	1.2	1.6
Dy	0.002	0.06	0.65	1.1	2.3
Er	0.004	0.10	0.7	1.0	3.2
Yb	0.008	0.11	0.7	0.9	4.0
Cr	0.2	2.0	10	12	2.0
Ni	10	4.0	2.0	3.0	0.4

eclogites^{12,14} and diamond inclusions^{13,14}. The existing data consistently point to ages >2,000 Myr for the establishment of the geochemical characteristics of the suites concerned. This old age, the geochemical paradox mentioned above and the apparently universal occurrence of this anomalous mantle under old continental crust have prompted the present investigation.

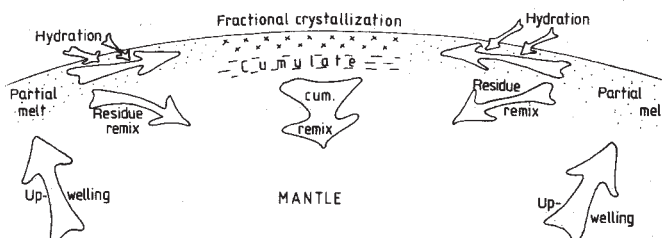
It has not been proved that the present-day mode of formation of continental crust as accretionary prisms and by igneous processes in and above subducted slabs goes back as far as the Archaean, and there are features which distinguish Archaean crust from late Proterozoic and Phanerozoic accreted terranes.

Table 2 Percentage fractionation for mode shown in Fig. 1

Mode	1	2	3
Olivine	15	15	5
Orthopyroxene	9	2	2
Clinopyroxene	1	8	8
Amphibole	20	20	25
Garnet	5	5	10
Total	50	50	50

In modern terranes, the volcanosedimentary sequences usually predominate over and predate tonalitic-granitic material. In contrast, in Archaean provinces granodioritic-tonalitic units are more abundant than greenstone belts, and granodioritic-tonalitic gneisses are normally the oldest rocks present. Furthermore, although greenstone belts in some instances resemble ophiolite complexes and may be at least in part tectonically emplaced¹⁵ they cannot generally be equated with ophiolites. They may be deposited on continental crust¹⁶, can exhibit platform facies and have never been observed to incorporate slices of mantle peridotite. Also, although there is a rough similarity in trace-element abundance patterns between Archaean tonalites^{17,18} and Phanerozoic calcalkaline volcanics and their derivatives^{19,20}, the former stand out by their large spread in heavy rare-earth element (HREE) abundances, extending to very low values (Fig. 1a) and by their strongly concave-up REE patterns. Ancient cratons are further characterized by the late Archaean intrusions of large masses of K-rich adamellites and granites which raise the average K and Rb content of the known continental crust to a level which is difficult to produce by modelling volcanic arcs²¹.

In the open-system fractional-crystallization model presented here, large near-surface fractionating magma reservoirs represent the continental crust, whereas the cumulate phases provide the material for the subcontinental mantle. The model is dynamic and continuous: the melt reservoirs are replenished by MORB-type magma, and the cumulate as well as the residue from MORB production are mixed back into the mantle after a delay (hence MANTle ReCYcling, MARCY, Fig. 2). This type of model follows naturally from the mantle convection modelling of Campbell and Jarvis²², which indicates that in the pre-Archaean an effectively worldwide shallow layer of partial melt existed, with maximum temperatures of ~1,580 °C being attained at a depth of ~30 km above sites of mantle upwelling (spreading centres). As long as this layer contains a large melt fraction it cannot be subducted, but the partial melt will move away from the loci of eruption and presumably gather in places farthest away from these sites. Initially this MORB of the pre-Archaean might be quite magnesian (almost komatiitic), but the thermal constraints²² would ensure that it would readily fractionate to an olivine tholeiitic composition. If there were an ocean a solidified hydrated crust would form quickly wherever magma reached the surface. This would be denser than the melt and therefore would frequently fall back into the magma causing it to become hydrous. The resulting cooling effect could be offset by the consequent lowering of the solidus. Gradual solidification would take the form of fractional crystallization and, because of the continuous addition of new melt, the behaviour would be similar to that described by O'Hara as open-system fractional crystallization²³. The amount of melt remaining in loci of such fractional crystallization and the major- and trace-element composition of this melt would approach a constant value when outputs (as cumulate) balance inputs (as melt). The major-element chemistry of the fractionating melt would thus stabilize to yield a cumulate of hydrated olivine tholeiite composition. Although the cumulate would probably differ modally from a tholeiite hydrated at low pressure, phase relations dictate that the melt would become similar to that produced by a small percentage partial melting of hydrated olivine-tholeiite—it could

**Fig. 2** Flow of matter in the MARCY model.

be roughly granodioritic-tonalitic²⁴.

With melt input and cumulate output, the open-system fractional crystallization model becomes an effective incompatible element pump. Ultimately, constant trace-element abundances of the fractionating melt body depend on the mineralogy of the steady-state cumulate. REE abundance patterns^{17,18} indicate that amphibole and garnet were in equilibrium with the liquid that formed the Archaean tonalites and granodiorites. Orthopyroxene and/or clinopyroxene as well as olivine are also likely crystallizing phases^{24,25}. Negative Eu anomalies are small or absent in the granodiorite-tonalite suite and even a very small amount of plagioclase fractionation would create a significant Eu anomaly. Thus, plagioclase would not have been an important fractionating phase: it either floated, or was converted to other phases at deeper levels in the magma body. Garnet in mafic rocks is generally associated with high pressures and its presence could therefore be used as an argument against a fractional crystallization model for early crustal genesis. However, the lower pressure boundary of the garnet field is highly dependent on the chemistry of the system, as can be seen by comparing metamorphic crustal rocks and peridotites. Nicholls²⁵ found garnet crystallizing from a hydrous andesitic melt at 15 kbar. Because the depth of the almost global super-solidus layer of Campbell and Jarvis²² extends well below 50 km, garnet fractionation must be considered a real possibility.

Some steady-state trace element concentrations for the fractionating melt, the cumulate and the mantle, calculated in a stepwise mode for various types of cumulates and using the partition coefficients listed in Table 1, are shown in Fig. 1b. The 'protocrust' REE patterns resemble those of the ancient tonalities (Fig. 1a) and are not very sensitive to variations in the relative abundance of cumulus phases. The uniformity of LREE abundances is especially striking. In models deriving tonalitic melts by partial melting of subducting slabs, such uniformity is difficult to achieve, as the LREE abundances would depend strongly on the percentage partial melt. Thus, the MARCY model can produce a realistic granodioritic-tonalitic protocrust in terms of both major and trace elements, but the K contents and Rb/Sr ratios are much too low for the average continental crust (Fig. 1a, b).

The steady-state cumulate, being essentially of tholeiitic composition, bears no resemblance to anomalous subcontinental mantle either in major- or in trace-element chemistry, although it may resemble some lower crustal nodules from Lesotho kimberlites²⁶. To transform this cumulate into anything resembling cold coarse peridotites¹ a depletion process involving extraction of melt as well as garnet has to be invoked. Such a process could follow naturally from the pressure-induced breakdown of amphibole. As the cumulate subsides, amphibole would breakdown at ~25 kbar (ref. 27), and more garnet and clinopyroxene would be formed, together with some hydrous melt. Even if the cumulate had by this time become completely solid, it would undergo remelting without a rise in temperature. The resulting suspension might facilitate the remixing of solid cumulate phases back into the mantle, but the hydrous melt produced would probably percolate upwards to higher levels of the cumulate or even into the fractionating melt body. Having equilibrated with a garnet-rich assemblage at the base of the cumulate, this

hydrous melt would be relatively depleted in HREE. It will probably also be rich in LILE, but could not produce anomalous subcontinental mantle by admixture to the cumulate: a depletion process is required involving more than just the base of the cumulate which is about to be mixed back into the mantle.

Such a process may be invoked in connection with tectonics following solidification of a portion of steady-state fractionating melt to crustal material. The moment that a sizeable block of granodioritic-tonalitic material solidified, the MARCY model would no longer apply to that section of the Earth. From this point in time, horizontal tectonics could leave their mark in the crustal shortening which is commonly observed in the Archaean^{15,16}. This has to accompany thickening of the crust and of the underlying cumulate. Such episodic thickening would cause amphibole breakdown on a much larger scale than during the preceding MARCY process: No new cumulate would be supplied, and the breakdown and consequent low-temperature remelting would occur wherever tectonics would force a portion of the cumulate through the amphibole high-pressure stability limit. Garnet and clinopyroxene (eclogite?) would be formed. The plasticity of the suspension could facilitate further crustal shortening, and tectonic kneading might lead to a density stratification of the reworked cumulate, with garnet especially settling at the bottom and being preferentially mixed back into the main body of the mantle. The hydrous melt generated in such episodes would scavenge K and incompatible trace elements from the cumulate during its generation and upward move, and might interact with the lower portion of the crust to cause extensive remelting. This mechanism for the generation of granites and adamellites in the late Archaean is complementary to the hypothesis of radioactive heating after crustal thickening²², and would help to explain the observation that these intrusives raise the K- and LIL trace-element content as well as the Rb/Sr ratio of the average crust considerably (Fig. 1c).

Whether this type of process would produce a realistic subcontinental mantle in major- and trace-element terms would depend on the amount of basaltic component depletion caused by the remelting episodes, and on the efficiency of gravity separation. These factors are poorly constrained, but I consider this a possible working hypothesis. Attempts at trace-element modelling have shown that many stages of amphibole breakdown, garnet settling and melt separation produce a stronger depletion in HREE than if the process occurs in a few stages. Fig. 1c shows that large-ion Lithophile (LIL) and REE patterns reminiscent of those in the subcontinental mantle can be produced. Furthermore, the Cr content becomes relatively high, which is in keeping with observations¹.

In the model, crustal shortening accompanied by thickening of the underlying cumulate and amphibole breakdown has to be the final act in the generation of large cratons stabilized by a thick rigid subcontinental mantle keel. As long as there is amphibole breakdown and the sub-continental mantle contains interstitial melt, it can be squeezed and crustal shortening is possible. When either all available amphibole has gone through the 25-kbar boundary, or all melt is squeezed out without further thickening, the subcontinental mantle keel loses its plasticity. This agrees with the observation that the intrusion of K-rich granites usually immediately precedes the stabilization of large cratonic blocks which are prevalent since the early Proterozoic. It is tempting to think that the progression from Archaean to Proterozoic tectonic styles is determined by the transition of the subcontinental mantle keel from a non-rigid to a rigid state.

I thank J. M. Barton, B. Harte, C. J. Hawkesworth, J. R. Ridley and J. F. Wilson for helpful comments.

Received 23 June; accepted 27 October 1986.

1. Harte, B. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C. J. & Norry, M. J.) 46-91 (Shiva, Orpington, 1983).
2. Nickel, K. G. & Green, D. H. *Earth planet. Sci. Lett.* 73, 158-170 (1985).
3. Harte, B. & Freer, R. *Terra Cognita* 2, 273-275 (1982).
4. Nixon, P. H., Rogers, N. W., Gibson, I. L. & Grey, A. A. *Rev. Earth planet. Sci.* 9, 285-309 (1981).
5. Frey, F. A. & Green, D. H. *Geochim. cosmochim. Acta* 38, 1023-1059 (1974).
6. Shimizu, N. *Phys. Chem. Earth* 9, 655-669 (1975).
7. Menzies, M. & Murthy, V. R. *Nature* 238, 634-636 (1980).
8. Smith, C. B. *Nature* 304, 51-54 (1983).
9. Hawkesworth, C. J., Rogers, N. W., van Calsteren, P. W. C. & Menzies, M. A. *Nature* 311, 331-335 (1984).
10. Kramers, J. D., Dawson, J. B. & Roddick, J. C. M. *Earth planet. Sci. Lett.* 65, 90-106 (1983).
11. Menzies, M. A. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C. J. & Norry, M. J.) 91-110 (Shiva, Orpington, 1983).
12. Jagoutz, E., Dawson, J. B., Hoernes, S., Spettel, B. B. & Wanke, H. *Lunar planet. Sci. Conf., Houston*, 395-396 (1984).
13. Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* 310, 198-202 (1984).
14. Kramers, J. D. *Earth planet. Sci. Lett.* 42, 58-70 (1979).
15. de Wit, M. J., Fripp, R. E. P. & Stanistreet, I. G. *Spec. Publ. geol. Soc. S. Afr.* 9, 21-29 (1983).
16. Wilson, J. F. *et al. Nature* 271, 23-27 (1978).
17. Arth, J. G. & Hanson, G. N. *Geochim. cosmochim. Acta* 39, 325-362 (1975).
18. Jahn, B. M., Glikson, A. Y., Peucat, J. J. & Hickman, A. H. *Geochim. cosmochim. Acta* 45, 1633-1652 (1981).
19. Hawkesworth, C. J., Hammill, M., Gledhill, A. R., van Calsteren, P. W. C. & Rogers, N. W. *Earth planet. Sci. Lett.* 58, 240-254 (1982).
20. Thirlwall, M. F. *Earth planet. Sci. Lett.* 58, 27-50 (1982).
21. Kay, R. W. *J. Geol.* 88, 497-522 (1980).
22. Campbell, I. H. & Jarvis, G. I. *Precamb. Res.* 26, 15-56 (1984).
23. O'Hara, M. J. *Nature* 266, 503-507 (1977).
24. Helz, R. T. *J. Petrol.* 17, 139-193 (1976).
25. Nicholls, I. A. *Contr. Miner. Petrol.* 45, 289-316 (1974).
26. Van Calsteren, P. W. C., Harris, N. B. W., Hawkesworth, C. J., Menzies, M. A. & Rogers, N. W. *J. Petrol.* (in the press).
27. Olafsson, M. & Eggler, D. H. *Earth planet. Sci. Lett.* 64, 305-315 (1983).
28. Wedepohl, K. H. *Forstsch. Miner.* 59, 203-205 (1981).
29. Shimizu, N. *Earth planet. Sci. Lett.* 25, 26-32 (1975).
30. Gurney, J. J., Harte, B. & Cox, K. G. *Phys. Chem. Earth* 9, 507-532 (1975).
31. Reid, A. M., Donaldson, C. H., Brown, R. W., Ridley, W. I. & Dawson, J. B. *Phys. Chem. Earth* 9, 525-557 (1975).
32. Irving, A. J. *Geochim. cosmochim. Acta* 42, 743-770 (1978).
33. Shimizu, N. *Geochim. cosmochim. Acta* 38, 1789-1798 (1974).
34. Stosch, H. G. *Geochim. cosmochim. Acta* 46, 793-811 (1982).

Provenance of Carboniferous sandstones from U-Pb dating of detrital zircons

Sarah Drewery, R. A. Cliff & M. R. Leeder

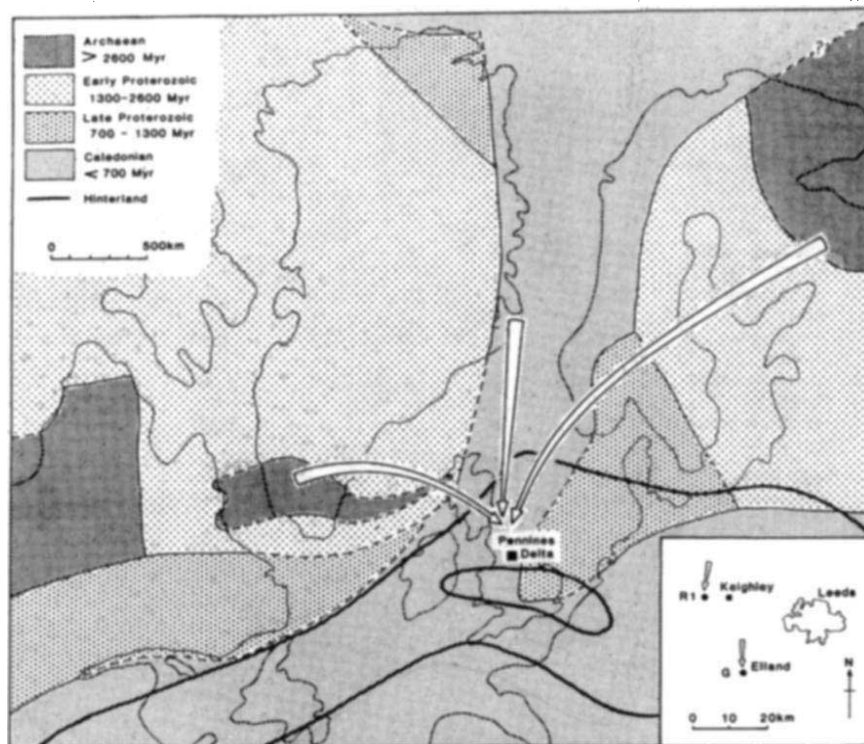
Department of Earth Sciences, The University, Leeds LS2 9JT, UK

One of the major problems of North Atlantic palaeogeographical development is the origin and source of the drainage system that deposited, during the Carboniferous, up to 5 km of clastic sediments across an east-west belt from western Ireland to western Poland. We present here results from U-Pb dating of detrital zircons in Namurian (mid-Carboniferous) fluvial clastics of northern England. These show a consistent and simple pattern with Carboniferous sandstones derived, in part, from active erosion of Archaean crystalline rocks possibly from western Greenland and/or Fennoscandia. Detritus recycled through the Caledonian is also represented but the proportion recycled through pre-existing sedimentary rocks, as implied by rounded zircon grains, is far smaller than expected.

Throughout the Carboniferous of northern Britain, enormous amounts of clastic sediment were supplied to extensional sedimentary basins by major river systems and their terminal deltaic complexes. Following Sorby's pioneering work¹ the northerly provenance of Namurian (Millstone Grit) sediments was determined by Gilligan's petrographic, heavy mineral and palaeocurrent studies². He deduced that the drainage basin covered a large part of Caledonian Scotto-Scandinavia, with the hinterlands being of granite-gneiss composition. Subsequent provenance studies^{3,4} have not seriously challenged these ideas. More recent facies studies have reinforced the suspicion that Namurian rivers were of huge magnitude. Evidence from giant cross-sets in Kinderscout Grit fluvial channel facies⁵ indicates contemporary channel depths of up to 35 m, implying that the channel complex was fed by a truly continental-scale drainage system.

The study of heavy minerals has long been recognized as an important tool in sedimentary provenance studies. Recently, isotopic dating of heavy minerals has been used to shed more light on the age structure of potential hinterlands indicated by

Fig. 1 Generalized Namurian depositional areas of the North Atlantic region showing major hinterland basement age provinces (simplified) and general location of source lands (including possible Archaean first-cycle detritus) suggested by palaeocurrent data. Inset: ■, sampling locations and palaeocurrent directions.



palaeocurrent and palaeogeographic reconstructions. Zircon and tourmaline have both been used in such studies⁶⁻⁹. Any heavy mineral population is likely to contain a mix derived from rock units of different ages, with the possibility of various stages of recycling. It is, therefore, essential to analyse separately small hand-picked samples of heavy minerals based on size, colour, morphology and magnetic susceptibility to minimize the risk of dating mixed populations.

The samples obtained for analysis were from the Kinderscout Grit group (Namurian stage, R1) and the Rough Rock (stage G) of the Central Pennine basin (Fig. 1). From the former group a 20-kg sample of Earl Crag Grit was collected from Hangingstone Quarry east of Cowling, Keighley Moor (SD 990430). McCabe⁷ interpreted the deposit as part of a distributary channel facies with a mean palaeocurrent direction roughly from the north (Fig. 1). The Rough Rock sample (15 kg) was taken from the base of a road cutting on the A629 at Elland near Halifax (SE 103213). This part of the sequence may represent a proximal mouth bar deposit (C. Bristow, personal communication); palaeocurrents are again from the north. Fine-grained sandstones were collected to facilitate comparisons between samples. Point-counting studies of each sample shows them to be sub-arkosic in composition. The stable heavy minerals zircon, tourmaline and rutile are present in both samples as are apatite and garnet and, interestingly, monazite which is usually considered diagnostic of granitic parent rocks.

Zircon fractions for analysis were hand picked from individual size fractions. Initial examination of the two zircon populations revealed in each two distinct colour groupings which are analogous to two zircon colour series of Gastil *et al.*¹⁰. (1) 'Colourless'-pale yellow-yellowish brown-pale brown. (2) 'Colourless'-pale pink-purple.

Colourless-pale brown grains are more common than pink-purple accounting for between 60% in the larger size fractions and 90% in the smaller size fractions. Each colour series includes grains of different morphologies, euhedral, sub-euhedral and rounded. Euhedral grains are more common in smaller size fractions and in the brown colour series but overall they are more abundant than the rounded grains. The euhedral grains may often be subdivided into two main groups implying tem-

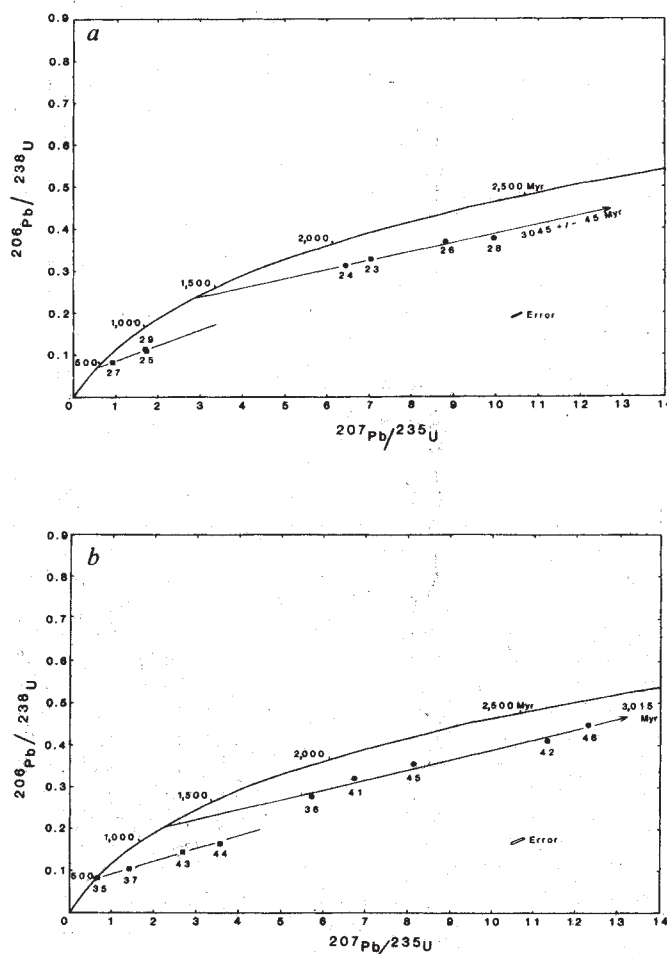


Fig. 2 Concordia diagrams for zircons: *a*, from Earl Crag sandstone (R1); *b*, from Rough Rock (G). ■, Colourless-brown zircon fractions; ●, purple zircon fractions.

Table 1 U-Pb data for Namurian samples

Sample	Size fraction (μm)	Mass (mg)	Concentrations		Atomic ratios				Apparent ages (Myr)		
			^{238}U (p.p.m.)	^{206}Pb (p.p.m.)	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{206}\text{Pb}$	$^{207}\text{Pb}/^{235}\text{U}$	$^{206}\text{Pb}/^{238}\text{U}$	$^{207}\text{Pb}/^{206}\text{Pb}$	$^{207}\text{Pb}/^{235}\text{U}$	$^{206}\text{Pb}/^{238}\text{U}$
R1 Earl Crag Grit											
27 Colourless, euhedral type I, NM1°*	<75	1.04	387	28	2,924	0.0826	0.940	0.0835	1,260	673	517
25 Brown, euhedral type I, NM1°	125-105	1.98	304	29	2,053	0.1147	1.729	0.1094	1,875	1,019	669
29 Brown, euhedral type I, NM1°	150-125	1.38	294	29	3,438	0.1083	1.702	0.1142	1,771	1,009	697
24 Purple, rounded, NM1°	105-75	1.22	272	74	3,600	0.1501	6.478	0.3132	2,347	2,043	1,756
23 Purple, rounded, NM1°	125-105	1.71	247	70	3,846	0.1563	7.059	0.3271	2,416	2,119	1,824
26 Purple, rounded, NM1°	150-125	1.37	210	68	6,068	0.1729	8.831	0.3700	2,589	2,321	2,029
28 Purple, euhedral type I, NM1°	125-105	1.10	335	110	3,321	0.1916	9.962	0.3774	2,756	2,431	2,064
G Rough Rock											
37 Colourless, sub-euhedral, NM2°	150-125	1.02	237	22	857	0.1001	1.444	0.1049	1,626	907	643
35 Brown, euhedral type I, NM1°	125-105	1.76	139	11	270	0.0603	0.676	0.0830	614	525	514
43 Brown, euhedral type I, NM2°	210-150	1.12	288	36	3,820	0.1352	2.691	0.1449	2,167	1,326	872
44 Brown, euhedral type II, NM2°	210-150	0.74	357	52	2,270	0.1578	3.571	0.1656	2,432	1,543	988
36 Purple, rounded, NM2°	105-75	1.54	269	69	253	0.1489	5.743	0.2760	2,333	1,938	1,572
41 Purple, rounded, NM2°	125-105	1.90	203	57	1,432	0.1523	6.765	0.3226	2,372	2,081	1,802
45 Purple, rounded, NM2°	210-150	1.03	177	55	5,590	0.1658	8.176	0.3580	2,516	2,251	1,973
46 Purple, rounded, NM2°	250-210	0.39	117	46	3,954	0.1981	12.30	0.4510	2,811	2,628	2,400
42 Purple, euhedral type I, NM2°	125-105	0.45	243	87	5,942	0.1997	11.33	0.4125	2,824	2,551	2,226

Method for U/Pb separation using anion exchange resin columns after ref. 27. R1 fractions spiked with a single combined ^{202}Pb - ^{236}U spike and G fractions with a ^{202}Pb - ^{235}U spike. Isotopic compositions measured using a VG Micromass 30 mass spectrometer. Pb isotopic fractionation monitored by frequent measurements of radiogenic lead standard SRM 983 (mean value $^{207}\text{Pb}/^{206}\text{Pb} = 0.07115 \pm 0.00008$) and data corrected as far as possible by normalizing to SMR = 0.07120. Apparent ages calculated with following constants: $^{235}\text{U}/^{238}\text{U} = 1/137.88$; $\lambda^{238}\text{U} = 1.55125 \times 10^{-10} \text{ yr}^{-1}$; $\lambda^{235}\text{U} = 9.8485 \times 10^{-10} \text{ yr}^{-1}$.

* NM, Non-magnetic at 1.45 A (Frantz magnetic separator); tilt is indicated in degrees. NM°, Non-magnetic at 2 A.

peratures of crystallization of 750 and 900 °C according to Pupin's typological classification¹¹.

Analytical data are given in Table 1 and are plotted on concordia plots in Fig. 2. All of the zircon fractions are discordant, although one fraction (G/35) of brown euhedral zircons from the Rough Rock plots very close to the concordia at 500 Myr (Fig. 2b). To interpret the pattern of discordance the following points must be considered.

First, the zircon population may contain a mixture from different source rocks. This problem is minimized by hand picking fractions for separate analysis.

Second, zircons from a single igneous source may not be homogeneous because of the frequent presence of inherited zircons. For example, Pidgeon and Aftalion¹² have shown that zircons from Caledonian granites contain inherited Precambrian components which may sometimes be seen optically as cores of 'old' zircon on which new zircon has grown.

Third, one or more episodes of lead-loss may have affected the zircons.

Purple zircon fractions in both samples are the oldest, with $^{207}\text{Pb}/^{206}\text{Pb}$ ages ranging from 2,330 to 2,820 Myr. Note that euhedral as well as rounded purple zircons produce Precambrian ages. The euhedral shape implies limited exposure to abrasion which suggests that the grains are first cycle, arguing for a Precambrian source directly feeding the basin. Conversely, rounded purple grains could have been transported huge distances through several sedimentary cycles because zircon is extremely resistant to destruction by abrasion. Purple zircons have long been associated with Precambrian rocks, being found in the Lewisian of north-west Scotland and in British sedimentary rocks from late-Precambrian to Jurassic in age¹³.

Figure 2 shows a general alignment of purple zircon fractions. This could be interpreted in terms of a single major Pb-loss/U-gain episode in each case; a best-fit line through those fractions from the Earl Crag Grit sample (Fig. 2a) has an upper intercept at $3,050 \pm 50$ Myr and a lower intercept at $1,360 \pm 70$ Myr. For the Rough Rock sample (Fig. 2b) the respective ages are $3,020 \pm 40$ Myr and $1,190 \pm 80$ Myr. In this model, the upper intercept represents the age of crystallization and the lower intercept the age of episodic Pb-loss. Minimum ages, given by $^{207}\text{Pb}/^{206}\text{Pb}$, are 2,760 Myr (Earl Crag Grit) and 2,820 Myr (Rough Rock); an initial age of crystallization of around 3,000 Myr for each is, therefore, not inconceivable. The lower intercept ages of 1,400 and 1,200 Myr, however, do not correspond to any well recorded events affecting Archaean rocks; this suggests that a more complex model may be required. There may have been more than one episodic Pb-loss. A two-stage model with Pb-loss events during the Laxfordian (1,700-1,800 Myr) and Caledonian (400-500 Myr) could explain the discordance in each sample. Providing that lead-loss during successive events was approximately proportional in the different fractions the upper intercept in such a model will still approximate the age of crystallization, even though the lower intercept has no significance.

The crystallization age of the purple-coloured zircons of ~3,000 Myr is older than that of most Lewisian zircons; ages between 2,660 and 2,800 Myr are common^{14,15} and are interpreted to date the Scourian granulite facies metamorphism. Zircons from quartzite lenses on Rona, however, have been interpreted by Bowes *et al.*¹⁶ in terms of a >3,200-Myr age ($^{207}\text{Pb}/^{206}\text{Pb}$ age = 2,840 Myr). Furthermore, Whitehouse and Moor bath¹⁷ report a Pb-Pb whole-rock age of $2,950 \pm 70$ Myr on amphibolite facies grey gneisses from the Lewisian. They

suggest that the preservation of this older age may indicate that whole-rock systems remain closed to U-Pb during amphibolite facies metamorphism. That is, Lewisian zircons that have only undergone amphibolite facies metamorphism may retain ages older than the date from the Scourian peak of granulite facies metamorphism.

The well-rounded shape of many of the purple grains and their frosted, pitted surfaces may imply a more distant source. The Nûk gneisses from the Godthaabsfjord area, western Greenland, for instance, produce primary ages of 2,890–3,065 Myr (ref. 18), and hyacinth zircons from the Amitsôq gneisses in the same region have been dated at $3,650 \pm 50$ Myr (ref. 19). What the exact sources may be, however, is impossible to tell without, at least, further knowledge of the extent of recycling.

The brown and colourless zircon fractions (mostly euhedral) have $^{207}\text{Pb}/^{206}\text{Pb}$ ages between 610 and 2,430 Myr. If a 'best-fit' line is fitted to the colourless-brown zircon fractions from the Rough Rock (Fig. 2b), the upper intercept age produced is $3,010 \pm 70$ Myr and the lower intercept 500 Myr. Ages of $2,730 \pm 170$ Myr and 420 ± 5 Myr respectively are obtained for the Earl Crag Grit zircons (Fig. 2a) although these are not very reliable because R1/25 and R1/29 plot so close together that the chord is essentially of two points only. The data could be a result of a failure to separate younger Caledonian zircons and older Archaean zircons, although a greater spread of data might be expected in this case. Alternatively, it is more likely they represent a reverse discordia pattern similar to those described by Pidgeon and Aftalion¹² for zircons from Caledonian granites in northern Britain. That is, the zircons are of Caledonian age but with an older, inherited Precambrian component. The 420-Myr age of the Earl Crag Grit zircons is (within error) consistent with the age of the voluminous suite of late-Silurian to early-Devonian granites exposed in the Highlands and Southern Uplands. The 500-Myr zircons from the Rough Rock are older, corresponding to a phase of Ordovician magmatism. These earlier granites were also exposed, however, as evidenced by clasts from Ordovician and early-Devonian conglomerates dated by Longman *et al.*²⁰ and by van Breemen and Bluck²¹. Zircon fraction G/35 (Fig. 2b) is very nearly concordant at 500 Myr indicating a minimal Pb-loss since formation. Fractions G/43 and G/44, however, appear to have a much larger component of Precambrian age (whether this is in the form of 'composite' grains or a mixture of grains of different ages).

Unpublished work on zircon fractions from other Carboniferous sandstones ranging in age from Dinantian to Westphalian has yielded similar results with 'old' purple zircons occurring throughout the Carboniferous. An earlier study, based on Sm-Nd whole-rock isotope data²², suggested that the sediments of southern Britain were formed by recycling of pre-existing sediments. This idea is contradicted by the new data presented here, as well as by subsequent Sm-Nd analyses²³.

The simplicity of these results, considering the continental-scale of the river system, is surprising. Mid-Proterozoic ages, for instance, do not seem to be represented in the detritus. This excludes areas such as the Ketilidian and Nagssugtoqidian of Greenland; van Breemen *et al.*²⁴ dated zircons from granitic belts from the Ketilidian of S. Greenland and found no evidence to support the theory of derivation from older basement. The Nagssugtoqidian of western Greenland is unsuitable because it has new mid-Proterozoic zircons in addition to those inherited from Archaean rocks. Similarly, the Laxfordian of north-west Scotland contains some new zircons²⁵. The proportion of detritus that has been recycled through pre-existing sedimentary rocks is not as great as anticipated; only ~10% of the zircon grains overall are well-rounded.

This evidence for sustained, active erosion of Precambrian basement in the late Palaeozoic presents a major sedimentary problem in that hypotheses for the development of the extensional suite of Carboniferous sedimentary basins²⁶ must also be accompanied by a theory to explain the establishment

of a huge continental-scale drainage system. The progradation of this depositional system into northern and central Britain reached its climax in Namurian times, with thick sequences of coarse feldspathic detritus being deposited. The persistence of first-cycle Archaean zircons throughout the Carboniferous implies that the drainage system was not altered substantially in areal extent in the Namurian, rather that a major increase of fluvial discharge and a reduction in hinterland chemical weathering occurred. The disappearance of semi-arid indicators in depositional facies in Namurian times, together with palaeomagnetic indicators of latitudinal shift may favour a climatic change (from semi-arid to monsoonal?) over a tectonic cause, but both variables may have been involved.

We thank members of the Department of Earth Sciences at Leeds for useful discussion. S.D. acknowledges receipt of a NERC studentship. Isotope work at Leeds is supported by NERC.

Received 21 July; accepted 22 October 1986.

1. Sorby, H. C. *Proc. Yorks. Geol. Polytech. Soc.* 3, 669–675 (1859).
2. Gilligan, A. Q. *Jl geol. Soc. Lond.* 75, 251–293 (1920).
3. Walker, C. T. thesis, Univ. Leeds (1952).
4. Muir, R. O. *Trans. Edinb. geol. Soc.* 19, 439–485 (1963).
5. McCabe, P. J. *Sedimentology* 24, 271–290 (1977).
6. Gaudette, H. E., Vitrac-Michard, A. & Allégre, C. J. *Earth planet. Sci. Lett.* 54, 248–260 (1981).
7. Gebauer, D. & Grunefelder, M. *Contr. Miner. Petrol.* 65, 29–37 (1977).
8. Grauert, B. *Contr. Miner. Petrol.* 40, 105–130 (1973).
9. Allen, P. *Proc. geol. Ass.* 86, 389–437 (1975).
10. Gastil, R. G., DeLisle, M. & Morgan, J. R. *Bull. geol. Soc. Am.* 78, 879–906 (1967).
11. Pupin, J. P. *Contr. Miner. Petrol.* 73, 207–220 (1980).
12. Pidgeon, R. T. & Aftalion, M. in *Crustal Evolution in Northwestern Britain and Adjacent Regions* (eds Bowes, D. R. & Leake, B. E.) 183–220 (Seel House Press, Liverpool, 1978).
13. Boswell, P. G. H. *Miner. Mag.* 21, 310–317 (1927).
14. Pidgeon, R. T. & Aftalion, M. *Earth planet. Sci. Lett.* 17, 269–274 (1972).
15. Pidgeon, R. T. & Bowes, D. R. *Geol. Mag.* 109, 247–258 (1972).
16. Bowes, D. R., Hopgood, A. M. & Pidgeon, R. T. *Geol. Mag.* 113, 545–552 (1976).
17. Whitehouse, M. J. & Moorbath, S. *Nature* 319, 488–489 (1986).
18. Baadsgaard, H. & McGregor, V. R. *Geochim. cosmochim. Acta* 45, 1099–1109 (1981).
19. Baadsgaard, H. *Earth planet. Sci. Lett.* 19, 22–28 (1973).
20. Longman, C. D., Bluck, B. J. & van Breemen, O. *Nature* 280, 578–581 (1979).
21. van Breemen, O. & Bluck, B. J. *Nature* 291, 113–117 (1981).
22. Miller, R. G. & O'Nions, R. K. *Earth planet. Sci. Lett.* 68, 459–470 (1984).
23. Davies, G., Gledhill, A. & Hawkesworth, C. *Earth planet. Sci. Lett.* 75, 1–12 (1985).
24. van Breemen, O., Aftalion, M. & Allaart, J. H. *Bull. geol. Soc. Am.* 85, 403–412 (1974).
25. van Breemen, O., Aftalion, M. & Pidgeon, R. T. *Scott. J. Geol.* 7, 139–152 (1971).
26. Leeder, M. R. Q. *Jl geol. Soc. Lond.* 139, 479–491 (1982).
27. Krogh, T. E. *Geochim. cosmochim. Acta* 37, 485–494 (1973).

Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere

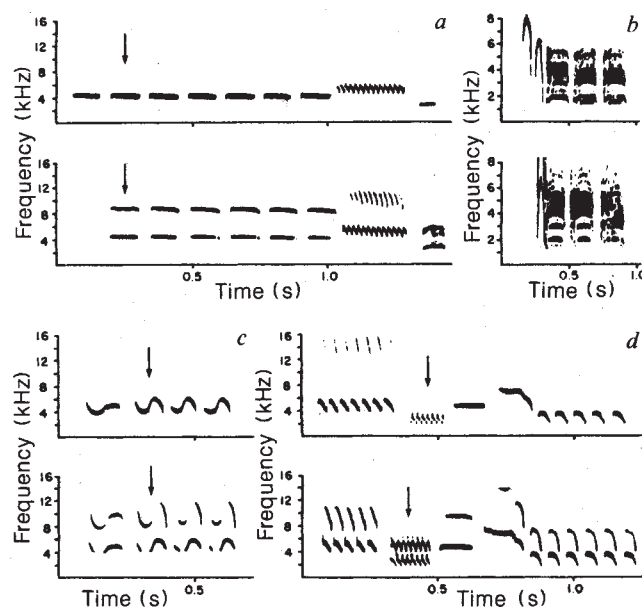
Stephen Nowicki

The Rockefeller University, Field Research Center, Millbrook, New York 12545, USA

The complexity and dependence on learning of many bird sounds have suggested parallels between birdsong and human speech^{1–4}, but the mechanisms by which each is produced have been supposed to differ markedly. In human speech, resonances of the vocal tract are thought to modulate in complex ways the sound produced by vibration of the vocal folds^{5–7}. The current theory of birdsong production holds that all variation in sound quality arises from the primary sound-producing organ, the syrinx, and that resonances of the vocal tract play no part^{8,9}. Here I present evidence, obtained from acoustic analyses of birdsongs recorded in a helium atmosphere, which contradicts this hypothesis. Not only does the songbird's vocal tract act as an acoustic filter, but its filter characteristics are actively coordinated with the output of the syrinx. Songbird and human phonation are thus more analogous than previously thought, in that both require coordination of an array of diverse motor systems.

One factor determining the acoustic resonances of a tube such as the vocal tract is sound velocity¹⁰. When atmospheric nitrogen

Fig. 1 Sonagrams of bird vocalizations produced in normal air (upper traces) and in the helium atmosphere (lower traces). *a*, *c*, *d*, Three different song sparrow (*Melospiza melodia*) songs (an entire song is depicted in *a*; selected song segments are shown in *c* and *d*). In helium there is a harmonic for every acoustic element normally appearing as a pure tone. *b*, Black-capped chickadee (*Parus atricapillus*) call. The higher frequency components of the terminal syllables are emphasized more in helium than in normal air. Arrows in *a* and *c* indicate syllables represented as amplitude spectra in Figs 2*a* and 2*b*, respectively. Arrow in *d* indicates a syllable for which the second harmonic, apparent in helium, occurs in the same frequency range as the fundamentals of adjacent syllables (Fig. 3). All birds were recorded in a 90 × 45 × 45 cm gas-sealed acoustic chamber, using a B&K 4145 calibrated microphone, a B&K 2604 microphone amplifier (flat weighting), and a Nagra 4.2L tape recorder (19 cm s⁻¹). After control vocalizations were recorded in normal atmosphere (upper traces), the chamber was flushed with 80% helium and 20% oxygen, and the bird recorded again (lower traces). Changes in tube resonances were monitored with a toy organ pipe inside the chamber, which showed a 67% increase in its resonance frequencies in heliox, corresponding to an increase in sound velocity from 331.5 m s⁻¹ to 554.9 m s⁻¹. Slight variations in timing and syllable composition in all vocalizations pictured are not a consequence of the helium atmosphere. Sonagrams were produced on a Kay Elemetric Digital Sona-graph (Model 7800); *a*, *c* and *d*, 16 kHz analysis range, 300 Hz filter bandwidth; *b*, 8 kHz analysis range, 300 Hz filter bandwidth. Other species so tested, but not pictured, are: swamp sparrow (*Melospiza georgiana*), starling (*Sturnus vulgaris*), chestnut-sided warbler (*Dendroica pensylvanica*), rufous-sided towhee (*Pipilo erythrophthalmus*), zebra finch (*Poephila guttata*), society finch (*Lonchura striata*), and brown-headed cowbird (*Molothrus ater*).



(comprising 80% of ordinary air) is replaced with the less dense gas helium, the velocity of sound increases from 331 m sec⁻¹ to 578 m sec⁻¹, and the resonance frequencies of a simple tube should increase by 74%^{11,12}. The effect of such an atmosphere on the singing of a bird depends on the manner in which resonances interact with the acoustic source. This point is best illustrated by two divergent examples. When a human speaks in helium, the fundamental frequency of vocal fold vibration does not change appreciably^{7,13,14}. The increase in pitch which is perceived is due to an upward shifting of the vocal tract resonances, or formants, with resulting emphasis on higher frequency overtones. In contrast, if a wind instrument like a trombone is played in helium, the fundamental frequency of the tone played does increase, in the same proportion as that calculated for resonances of the corresponding tube^{15,16}. Acoustic resonances modify the sound output in each case; the difference is due to the degree to which the sound source is acoustically coupled to the resonator. In the trombone, the resonator is strongly coupled to the source, largely determining its mode of vibration, whereas the resonator acts predominantly as a filter in the human voice.

The fact that the two sides of an oscine bird's syrinx seem able to operate independently originally led to the hypothesis that vocal tract resonances play no role in bird phonation⁸. Two harmonically unrelated tones often occur simultaneously in birdsong, which was thought to be impossible if the trachea acted as a coupled resonator. Nor was the vocal tract presumed to act as an uncoupled acoustic filter, since no relationship could be found between the amplitude and frequency of song syllables and predicted tracheal resonances⁸. This second argument is less compelling since it is assumed that filter properties must be static (in contrast to humans where resonances actively change during speech).

I recorded the songs and calls of nine species of oscine birds both in normal air and in a helium atmosphere. The results are equivalent for all species, and are typified by several song types of the song sparrow (*Melospiza melodia*) (Fig. 1*a*, *c*, *d*) and the call of the black-capped chickadee (*Parus atricapillus*) (Fig. 1*b*). The most striking effect of helium is that, to each song element

appearing as a pure tone in normal atmosphere, a harmonic overtone is added (Figs 1*a*, 2*a*). The fundamental frequency of these tones increases in helium by 3–5% (Fig. 2*a*), an order of magnitude less than the 67% shift that is observed in the toy organ pipe used for calibration. These notes also decrease in amplitude, often by over 10 decibels (dB). Sounds with a broad band of frequencies, such as the terminal syllables of the chick-

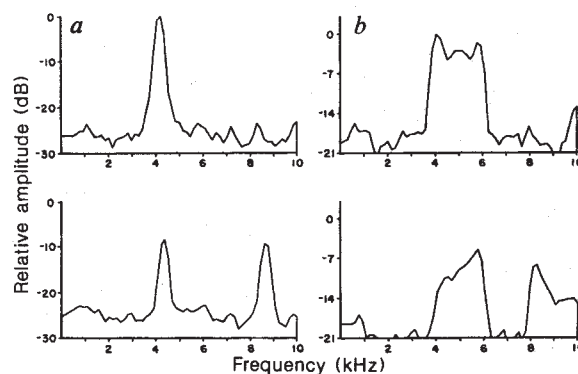


Fig. 2 Amplitude spectra of syllables produced in normal air (upper traces) and in the helium atmosphere (lower traces). Each spectrum is an average calculated over the entire duration of a single syllable (*a* corresponds to arrows in Fig. 1*a*; *b* corresponds to arrows in Fig. 1*c*). The amplitudes of signals in helium have been increased in gain by 10 dB compared to their counterparts in normal air to illustrate optimally the emergent frequency components. Note that if a vocal filter is centred at 4,200 Hz, the frequency of the syllable illustrated in *a* (and a realistic resonance value given the dimensions of the bird's vocal tract), then its centre frequency would shift to 7,000 Hz in the helium atmosphere, a value intermediate between the fundamental frequency and its harmonic. The small shift observed in the fundamental frequencies is similar to that observed in human speech^{13,14}, and may be caused by differential loading on the vibrating structure in the less dense atmosphere. Spectra were generated digitally, using a 128 point DFT, 10 kHz analysis range, and 156 Hz frequency resolution.

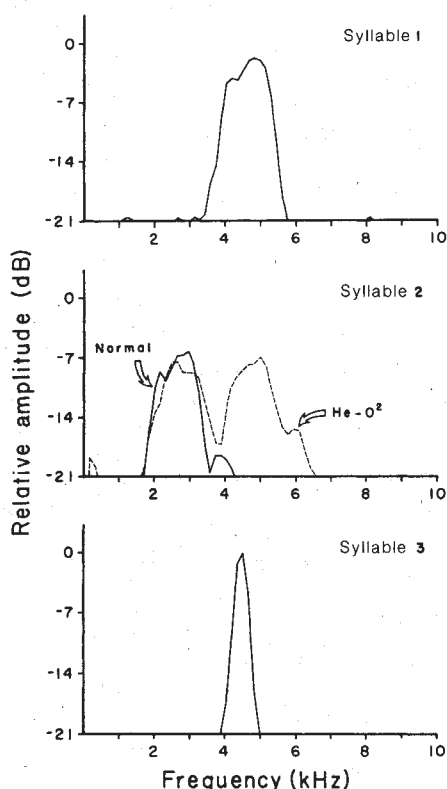


Fig. 3 Amplitude spectra of three consecutive syllables in normal atmosphere (solid lines), illustrating the overlap in frequency ranges of the second harmonic component of syllable 2, apparent only in helium (dashed line), and the normal fundamental frequency components of syllables 1 and 3. A change in the frequency response characteristics of the vocal filter must occur from syllable to syllable in order to account for the lack of energy in the second harmonic component of syllable 2 in normal atmosphere. Each average spectrum is calculated digitally, as in Fig. 2, over the entire duration of a syllable. Syllable 2 corresponds to the rapidly modulated syllable marked with arrows in Fig. 1d, while syllables 1 and 3 correspond to the trilled syllable immediately preceding and the tonal syllable immediately following, respectively. The amplitude spectrum in helium (dashed line) is shown for syllable 2 only.

adee's call, display a shift in the amplitudes of frequency components, with higher frequencies emphasized in helium (see Fig. 1b).

The changes observed in both pure tone and broad-band sounds are explained if the bird's vocal tract is acting as an acoustic band-pass filter that is not significantly coupled to the vibration source. Thus, the shifted emphasis to higher frequencies in the chickadee's broad-band syllable (Fig. 1b) is analogous to the change observed for a human voice in helium. The pure tone case (Figs 1a, 2a) indicates that the 'whistles' common in birdsong are not pure tones, but rather harmonic sounds with relatively high frequency fundamentals. Normally the energy in their second harmonics is strongly attenuated by a filter centred around the fundamental frequency. In helium, this filter is shifted upward, revealing the harmonic content. Since neither the fundamental frequency nor the second harmonic is now squarely centred on the filter's best frequency, both have reduced amplitudes compared to the fundamental in ordinary atmosphere. In a widely cited earlier study, Hersh¹⁷ recorded two species in helium-air, both of which sang 'whistled' songs with fundamental frequencies above 4 kHz. Because he could only analyse sounds up to 8 kHz (a typical upper frequency limit on

older spectrographs), Hersh failed to detect the emergence of harmonics and mistakenly concluded that no change occurred.

Further evidence for a vocal filter comes from frequency modulated whistles, the averaged spectra of which have broad frequency distributions with two distinct amplitude peaks (Figs 1c, 2b). In helium, the lower peak of the fundamental component is significantly more attenuated than the higher one. The reverse is true of the second harmonic component. These amplitudes can only be accounted for by shifting the centre frequency of a band-pass filter between the two harmonic components; the sloping shoulders of the filter function result in greater attenuation of the peaks further from the centre frequency.

As a bird utters a complex song, the centre frequency of its vocal filter must track actively the fundamental frequency of each syllable. Evidence of tracking comes from cases in which the second harmonic of a syllable falls in the same frequency range as the fundamental of an adjacent syllable (Fig. 1d). This harmonic would appear in normal atmosphere if the frequency response of the filter were invariant from syllable to syllable (Fig. 3). Filtering out of the harmonic, which clearly occurs, requires rapid shifting of the centre frequency of the filter before production. A bird could adjust its vocal filter in several ways: for example by varying tracheal length, by constriction of the larynx, or by flaring its throat and beak. Such configurational changes could well correspond to the head movements commonly observed in singing birds.

The parallels drawn between birdsong and human speech¹⁻⁴ can now be extended to the level of production. Like human speech, and in contrast to previous theories, birdsong must be viewed as the coordinated output of several motor systems acting in concert. There are contrasts as well, perhaps the greatest being that the oscine bird possesses two potentially independent sound sources. If these sources separately produce sounds, both must fall within the band-pass frequency range of the vocal filter. Further, oscine sounds have fundamentals on the order of several thousand Hz as opposed to human speech fundamentals usually less than 200 Hz. The bird's vocal filter tends to concentrate energy at a single frequency because of the wide separation of harmonic components, whereas the human vocal filter varies the distribution of energy within a wide band of frequencies. Lastly, the bird's filter probably simply tracks the output of the syrinx, whereas the more complex human vocal tract both modulates laryngeal output and generates its own phonemic gestures.

I thank R. R. Capranica, C. W. Clark, A. S. Gaunt, D. R. Griffin, S. Peters, K. N. Stevens, and especially P. Marler for discussion and comments on the manuscript. Supported by Biomedical Research Support Grant S07 RR07065 to The Rockefeller University and USPHS Post-doctoral Fellowship 5 F32 NS07520-02.

Received 7 May; accepted 20 October 1986.

1. Marler, P. *Am. Scient.* **58**, 669-673 (1970).
2. Marler, P. & Peters, S. in *Perspectives on the Study of Speech* (eds Eimas, P. D. & Miller, J. L.) 75-112 (Lawrence Erlbaum Ass., Hillsdale, New Jersey, 1981).
3. Nottebohm, F. *Prog. Psychobiol. & physiol. Psychol.* **9**, 85-124 (1980).
4. Konishi, M. *A. Rev. Neurosci.* **8**, 125-170 (1985).
5. Fant, G. *Acoustic Theory of Speech Production* (Mouton, The Hague, 1960).
6. Flanagan, J. L. *Speech Analysis, Synthesis, and Perception*, 2nd edn (Springer-Verlag, Berlin, 1972).
7. Lieberman, P. *Speech Physiology and Acoustic Phonetics* (Macmillan, New York, 1977).
8. Greenewalt, C. H. *Bird Song: Acoustics and Physiology*, 55-61 & 79-87 (Smithsonian, Washington, D.C., 1968).
9. Gaunt, A. S. & Gaunt, S. L. L. in *Current Ornithology*, Vol. 2 (ed. Johnston, R. F.) 213-246 (Plenum, New York, 1985).
10. Lord Rayleigh *The Theory of Sound* 2nd edn (Dover, New York, 1945).
11. Hardy, H. C., Telfair, D. & Pielemeier, W. H. *J. acoust. Soc. Am.* **13**, 226-233 (1942).
12. Beranek, L. L. *Acoustic Measurements*, 37-49 (Wiley, New York, 1949).
13. Beil, R. G. *J. acoust. Soc. Am.* **34**, 347-349 (1962).
14. Holywell, K. & Harvey, G. *J. acoust. Soc. Am.* **36**, 210-211 (1964).
15. Martin, D. W. *J. acoust. Soc. Am.* **13**, 305-308 (1942).
16. Benade, A. H. *Fundamentals of Musical Acoustics*, 391-429 (Oxford University Press, London, 1976).
17. Hersh, G. L. Thesis, Univ. California, Berkeley (1966).

Multiple classes of glutamate receptor on depolarizing bipolar cells in retina

S. Nawy & D. R. Copenhagen

Department of Physiology, Division of Neuroscience, University of California at San Francisco, San Francisco, California 94143, USA

Multiple subtypes of excitatory amino acid receptor have been found on individual dissociated neurones¹. These findings were obtained from cells without intact synaptic connections, so the functional roles for such receptor subtypes are unknown. We have recorded intracellular responses from depolarizing bipolar cells (DBC) that receive direct synaptic input from two distinct populations of neurones: rods and cones²⁻⁴. We report here that 2-amino-4-phosphonobutyrate (APB), a glutamate analogue, reveals two subtypes of glutamate receptors on DBCs. APB acts on the same receptor that mediates synaptic transmission from rods but has no action on the second subtype of glutamate receptor. These results show that the rod and cone inputs to DBCs are mediated by pharmacologically distinct receptors and that subtypes of glutamate receptor existing on single neurones can subserve separate, functionally defined synaptic inputs.

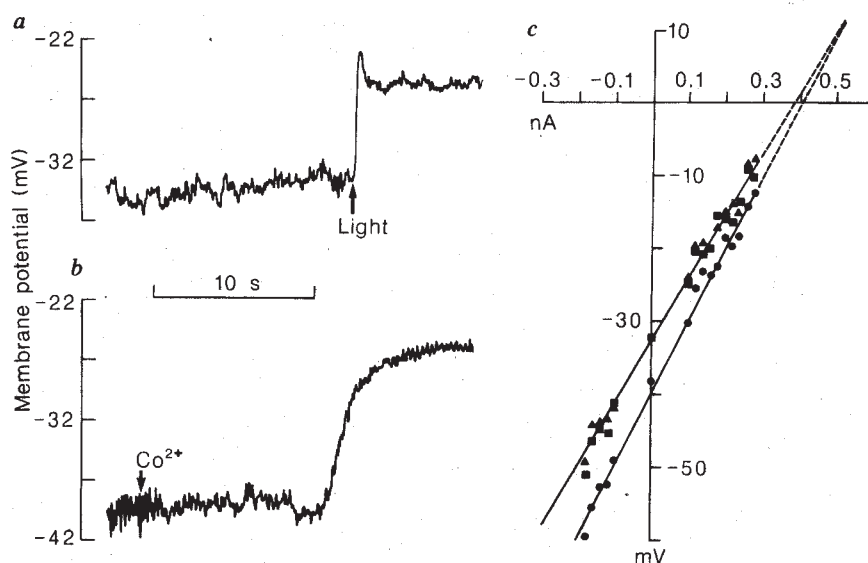
We examined the conductance mechanism underlying the rod-driven light response in depolarizing bipolar cells, using a technique described previously⁵ to eliminate cone input. Figure 1a shows the maximal response of a bipolar cell to a spot of light. Application of cobalt (Co^{2+}) to the superfusate depolarized the cell to the same potential as light (Fig. 1b). Previous studies in fish retina have shown that this concentration of Co^{2+} blocks transmitter release from photoreceptors⁶. Thus it appears that saturating intensities of light completely suppress transmit-

ter release from rods. To measure membrane conductance, we injected brief pulses of hyperpolarizing and depolarizing current of varying amplitude through the intracellular recording electrode. The resulting voltage change across the membrane was measured and plotted (see ref. 7). An example of such a current-voltage relation obtained from another bipolar cell under continuous illumination, Co^{2+} and in the dark is shown in Fig. 1c. The reduced slope of the light and Co^{2+} plots indicates an increased membrane conductance relative to darkness. The extrapolated reversal of the light and Co^{2+} responses was positive relative to resting potential. These current-voltage relations indicate that the rod transmitter, released continuously in the dark⁸ and blocked both by light and Co^{2+} , decreases membrane conductance by closing ionic channels whose reversal potential is more positive than the dark potential. This conductance mechanism has been proposed for the rod transmitter by Saito and colleagues^{4,9}, who demonstrated reversal of the rod light response at potentials positive to resting potential.

Conductance changes produced by light and the glutamate analogue APB were compared. APB, acting as an agonist, has been shown to block light responses in depolarizing bipolar cells in the all-rod retina of the dogfish¹⁰, and in the mixed rod-cone amphibian retina^{11,12}. Figure 2a shows the response of a bipolar cell to the superfusion of 2 μM APB. The membrane was hyperpolarized and the light responses were blocked. Current-voltage relations for this cell were measured in darkness, during continuous illumination and during superfusion of APB (Fig. 2b). The extrapolation of the current-voltage relations obtained in APB crosses the dark current-voltage plot at almost exactly the same potential as the plot for light. These data support the conclusion that APB closes the same channels that are closed by the native rod transmitter. The action is similar to that found in dogfish retina, contradicting the observation that APB has no effect on bipolar cells in the teleost retina¹³.

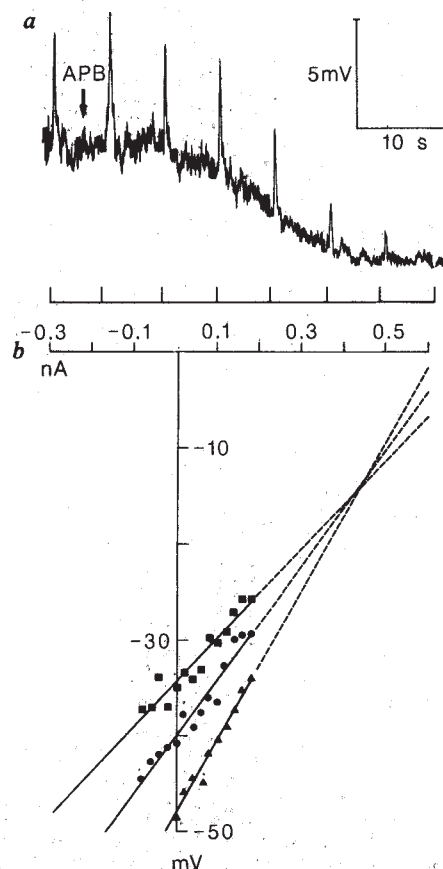
The conductance change underlying the action of glutamate,

Fig. 1 a, Responses of a depolarizing bipolar cell to a light of wavelength 550 nm and intensity of 3.9 photons $\text{sec}^{-1} \mu\text{m}^{-2}$; b, response of the same cell to 1 mM Co^{2+} added to the normal solution; c, current-voltage relations for a different cell obtained in the presence of $\blacktriangle$, 1 mM Co^{2+} ; $\blacksquare$, light of 1.96 photons $\text{sec}^{-1} \mu\text{m}^{-2}$; and $\bullet$, in darkness (control). These relations were obtained by injecting 500 millisecond pulses of current and measuring the average resulting steady state voltage deflection across a bridge circuit. The voltage step usually reached its final value within 100 milliseconds. Identical pulses of current were injected into the electrode after pulling out of the cell and the resulting voltage steps were then subtracted from the intracellularly recorded voltage steps in order to obtain the potential drop across the cell membrane. The lines are the best least-squares fit to the data. For clarity, no line is drawn for Co^{2+} . Slope of the current-voltage relations in dark, 99 M Ω ; light, 86 M Ω ; Co^{2+} (not shown), 87 M Ω .



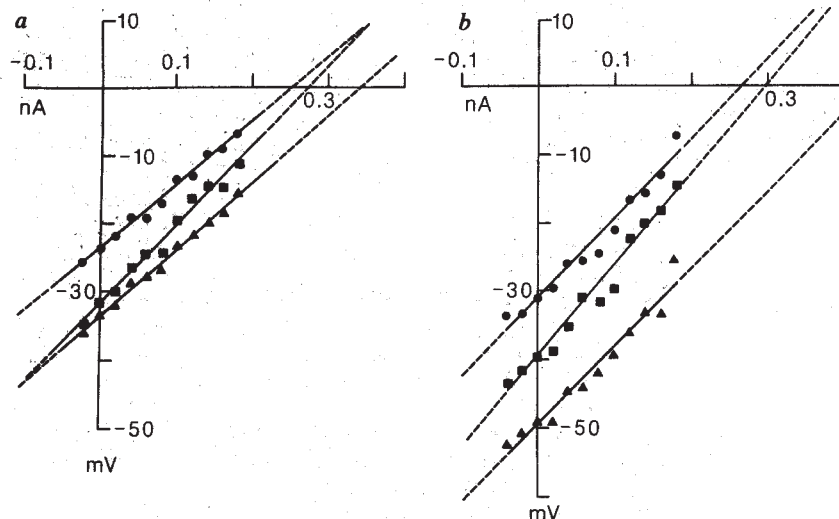
Methods. Current-voltage relations were measured in 24 depolarizing bipolar cells. Retinas from dark-adapted fish that were maintained in an outdoor pond were isolated from the pigment epithelium under infrared light. Fish were placed in darkness 3–6 h after light onset and kept there for at least 1 h before dissection. We have shown previously that neuronal responses from retinas isolated in this way are almost all rod driven⁵. Retinas were then incubated in darkness in a 20% solution of Wydase hyaluronidase in superfusion medium for 20–30 min at 4 °C and mounted photoreceptor-side up on a doughnut-shaped piece of filter paper in the recording chamber. We found that hyaluronidase effectively dissolved any vitreous remaining on the isolated retina, and allowed better adhesion of the retina to the filter paper. Mounted retinas were superfused continuously with oxygenated superfusion solution at room temperature (20–22 °C). This solution was L15 culture medium modified to contain: NaCl, 120 mM; MgSO_4 , 1.2 mM; KCl, 2.5 mM; CaCl_2 , 2.2 mM. The solution also contained 10 mM glucose and was buffered to pH 7.8 in 3 mM hydroxyethyl piperazine ethane sulphonic acid (HEPES). Depolarizing bipolar cells penetrated with glass micro-electrodes pulled on a Brown-Fleming puller¹⁶ and filled with 2 M potassium acetate. Electrode resistances varied from 500 to 800 M Ω .

Fig. 2 *a*, Response of a bipolar cell to APB. At the arrow, the normal solution was switched to one containing $2\ \mu\text{M}$ APB (D, L-APB, obtained from Sigma). The responses shown are to 10 millisecond flashes of light of 9.8×10^{-3} photons $\text{flash}^{-1}\ \mu\text{m}^{-2}$. The timing of the flashes is shown below. *b*, Current-voltage relations obtained from the same cell $\bullet$, in darkness; $\blacksquare$, in a steady background of 9.8 photons $\text{sec}^{-1}\ \mu\text{m}^{-2}$; and $\blacktriangle$, during superfusion of APB. The extrapolated reversal potentials of light and APB are nearly identical. Slope in darkness, 59 $\text{M}\Omega$; in light, 45 $\text{M}\Omega$; in APB, 77 $\text{M}\Omega$.



by contrast, was very different from the action of either the native rod transmitter or APB. In Fig. 3*a*, current-voltage relations obtained during separate superfusion of $10\ \mu\text{M}$ APB, 2 mM glutamate and control solution are shown for the same cell. All measurements were made in darkness. Once again, APB produced a hyperpolarization and decreased membrane conductance. Glutamate blocked the light response and produced an even larger hyperpolarization of the membrane than APB. But there was little or no apparent conductance change associated with the action of glutamate.

The discrepancy between the current-voltage relations in APB and glutamate can be explained most easily if glutamate is acting at both the APB-sensitive site and an additional type of site. Figure 3*b* shows an experiment in which $10\ \mu\text{M}$ APB was first perfused onto the retina. This concentration of APB was five-fold higher than necessary to block the light response. In the presence of APB, the retina was then superfused with 2 mM glutamate.



The slope of the current-voltage relation was decreased relative to APB alone. Thus in the presence of APB, the ability of glutamate to open channels with a negative reversal potential is unmasked. We therefore postulate that the apparent lack of a conductance change produced by glutamate is due to the summation of two simultaneous actions, the closing of APB-sensitive channels and the opening of APB-insensitive channels. Both actions of glutamate differ from those found at other excitatory amino acid receptors in the CNS^{1,14,15} which open channels with a reversal potential of about 0 mV.

Our results provide strong evidence that the rod and cone pathways use pharmacologically distinct receptors on the bipolar cell membrane. One receptor, mediating rod signals, is similar to the DBC receptor previously described in the retina¹⁰⁻¹²; it is sensitive to APB and acts by closing ionic channels. The second receptor is pharmacologically distinct from the first, as it is insensitive to APB and, unlike the first receptor, it

Fig. 3 *a*, Current-voltage relations obtained as described above in $\bullet$, control solution; $\blacksquare$, $10\ \mu\text{M}$ APB; and $\blacktriangle$, 2 mM glutamate (Sigma). Superfusion of APB produced a hyperpolarization of the membrane and an increase in slope. The reversal potential of the APB effect was positive to the dark potential. Superfusion of the cell with 2 mM glutamate also produced a hyperpolarization of the membrane, but a much smaller conductance change. Slope in the dark was 94 $\text{M}\Omega$; in APB, 117 $\text{M}\Omega$; in glutamate, 99 $\text{M}\Omega$. *b*, Current-voltage relations for a different cell during superfusion of: $\bullet$, control medium; $\blacksquare$, $10\ \mu\text{M}$ APB; and $\blacktriangle$, $10\ \mu\text{M}$ APB + 2 mM glutamate. Slope in darkness was 116 $\text{M}\Omega$; in APB, 134 $\text{M}\Omega$; in APB + GLU, 113 $\text{M}\Omega$.

opens ionic channels. The reversal potential of ion flow in these channels is similar to the reversal potential of the cone light response, as shown previously^{4,8}. Thus, this receptor probably mediates cone signals. We have described an example of a single neurone which has two types of glutamate receptor, each type receiving synaptic input from separate and distinctly identifiable presynaptic sources.

We thank Drs R. Nicoll, J. Belgium and G. Ayoub, and S. Mittman and L. Stone for a critical reading of the manuscript. This work was supported by NIH EY01869 (to D.R.C.) and NIH training grant EY07058.

Received 24 March; accepted 3 October 1986.

1. Mayer, M. L. & Westbrook, G. L. *J. Physiol., Lond.* **361**, 65-90 (1985).
2. Ishida, A. T., Stell, W. K. & Lightfoot, D. O. *J. comp. Neurol.* **191**, 315-335 (1980).
3. Kaneko, T. & Tachibana, M. *Sensory Processes* **2**, 383-387 (1978).
4. Saito, T., Kondo, H. & Toyoda, J. *Vis. Res.* **18**, 591-595 (1978).
5. Nawy, S. & Copenhagen, D. R. *Soc. Neurosci. Abstr.* **10**, 324 (1984).
6. Kaneko, A. & Shimazaki, H. *Cold Spring Harb. Symp. quant. Biol.* **40**, 537-546 (1976).
7. Belgium, J. H., Dvorak, D. R. & McReynolds, J. S. *J. Physiol., Lond.* **326**, 91-108 (1982).
8. Dowling, J. E. & Ripps, H. *Nature* **242**, 101-103 (1973).
9. Saito, T., Kondo, H. & Toyoda, J. *J. gen. Physiol.* **73**, 73-90 (1979).
10. Shiels, R. A., Falk, G. & Naghshineh, S. *Nature* **294**, 592-594 (1981).
11. Slaughter, M. M. & Miller, R. F. *Science* **211**, 182-185 (1981).
12. Slaughter, M. M. & Miller, R. F. *J. Neurosci.* **5**, 224-233 (1986).
13. Miller, R. F. & Slaughter, M. M. in *Retinal Transmitters and Modulators: Models for the Brain*, Vol. 2 (ed. Morgan, W. W.) 123-160 (CRC Press, 1985).
14. Watkins, J. C. & Evans, R. H. A. *Rev. Pharmac.* **21**, 165-205 (1981).
15. Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462-465 (1984).
16. Flaming, D. G. & Brown, K. T. *J. Neurosci. Meth.* **6**, 91-102 (1982).

Activation of protein kinase C augments evoked transmitter release

Roy Shapira*, Shai D. Silberberg*, Simona Ginsburg*† & Rami Rahamimoff*

* Department of Physiology, Hebrew University-Hadassah Medical School, PO Box 1172, Jerusalem 91010, Israel

† Everyman's University, PO Box 39328, Tel Aviv, Israel

In view of the emerging role of the phosphoinositide system in cellular communication^{1,2} we examined its involvement in quantal-transmitter release, which is a key element in synaptic transmission³. Transmitter release is normally activated by an increase in intracellular calcium, achieved either by entry of calcium ions through the presynaptic membrane^{4,5} or by intracellular calcium liberation⁶. One of the targets of the phosphoinositide signalling system is the enzyme protein kinase C (PKC), which can be activated experimentally by tumour promoting phorbol esters, including 12-O-tetradecanoylphorbol-13-acetate (TPA)⁷. Such activation of PKC may be implicated in transmitter release in two ways. First, phorbol esters were found to increase secretion⁸⁻¹³ and enhance calcium currents^{14,15}; it might therefore be expected that they would increase synaptic transmitter release. But phorbol esters also inhibit the calcium current in dorsal root ganglion neurones¹⁶. We report that the phorbol ester TPA augments synaptic transmission at the neuromuscular junction by increasing transmitter liberation. Activation of PKC also deepens synaptic depression.

Activation of the phosphoinositide pathway produces two internal messengers (inositol 1,4,5-triphosphate (IP₃) and diacylglycerol (DG)) which are formed by splitting phosphatidylinositol 4,5-bisphosphate. IP₃ is involved in liberation of calcium from internal stores, DG activates PKC, which, in its active form, is associated with the plasma membrane^{17,21} and endogenous proteins in neuronal cultures²² and *Aplysia* ganglia²³. Since PKC activation can be mediated directly by phorbol esters in the absence of phosphoinositide breakdown, we used TPA to examine the effects of such activation on synaptic transmission.

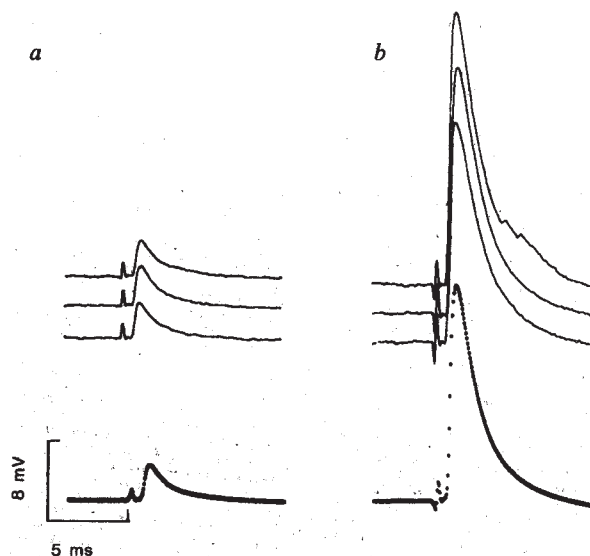


Fig. 1 The effect of TPA on evoked synaptic potentials. *a*, Individual (top) and average (bottom) e.p.s. recorded from a cell bathed in control Ringer solution (0.3 mM Ca + 1.0 mM Mg); *b*, Individual (top) and average (bottom) e.p.s. from the same cell as in *a*, 1 h after applying the test solution (162 nM TPA, 0.01% DMSO). Bottom traces represent averages of 470 individual responses. The experiment was performed at room temperature and the stimulation frequency was 0.33 Hz.

Most experiments were performed on the frog sartorius nerve-muscle preparation (except the experiments on depression, Fig. 3, which were done on the cutaneous pectoris). The preparation was mounted in a 7 ml bath and constantly perfused with control or test solutions at a rate of 1-2 ml min⁻¹. The experiments were performed at room temperature. End plate potentials (e.p.s), evoked by stimulating the nerve at 0.33 Hz, and miniature end plate potentials (m.e.p.s) were recorded intracellularly by conventional electrophysiological techniques and sampled by a PDP MINC 11/23 microcomputer. Synaptic potential amplitudes, quantal contents of e.p.s and m.e.p.p. frequencies were analysed by programmes designed for this purpose. Control solutions contained 116 mM NaCl, 2 mM KCl, 0.3 mM CaCl₂, 1-1.5 mM MgCl₂ and 5 mM hydroxyethyl piperazine ethane sulphonic acid (HEPES) buffer (pH 7.3). Test solutions contained either 162 nM (100 ng ml⁻¹) of the active phorbol-ester TPA, or 158 nM (100 ng ml⁻¹) of the analogue 4-O-M-TPA (phorbol 12-myristate 13-acetate 4-O-methyl ether). Both test solutions contained 0.01% dimethyl sulphoxide (DMSO), which at this concentration is unlikely to have any effect on transmitter release²⁴.

The main effect of TPA on evoked synaptic potentials is shown in Fig. 1. The average e.p.p. amplitude was significantly enhanced after applying the test solution (162 nM TPA). Similar results were obtained in 14 additional experiments, with an average enhancement in e.p.p. amplitude of 174% ($\pm 19.9\%$ s.e.m.). In 5 cases, TPA produced such a large increase in the synaptic response that muscle twitches terminated the experiments. The observed onset of TPA action varied considerably from 5 to 140 minutes in the different preparations. We found that the effect of TPA could not be reversed by up to a one-hour wash in control solution. TPA also caused a marked increase in the frequency of spontaneous release, with no substantial change in m.e.p.p. amplitude. In 5 experiments, the average enhancement in m.e.p.p. frequency, measured one hour after the onset of the effects, was found to be 395% ($\pm 111\%$ s.e.m.).

These effects of TPA can be demonstrated in a single cell (Fig. 2). The e.p.p. amplitudes increased after application of TPA (Fig. 2*a*), whereas m.e.p.p. amplitudes remained fairly constant throughout the experiments (Fig. 2*b*); hence, a marked

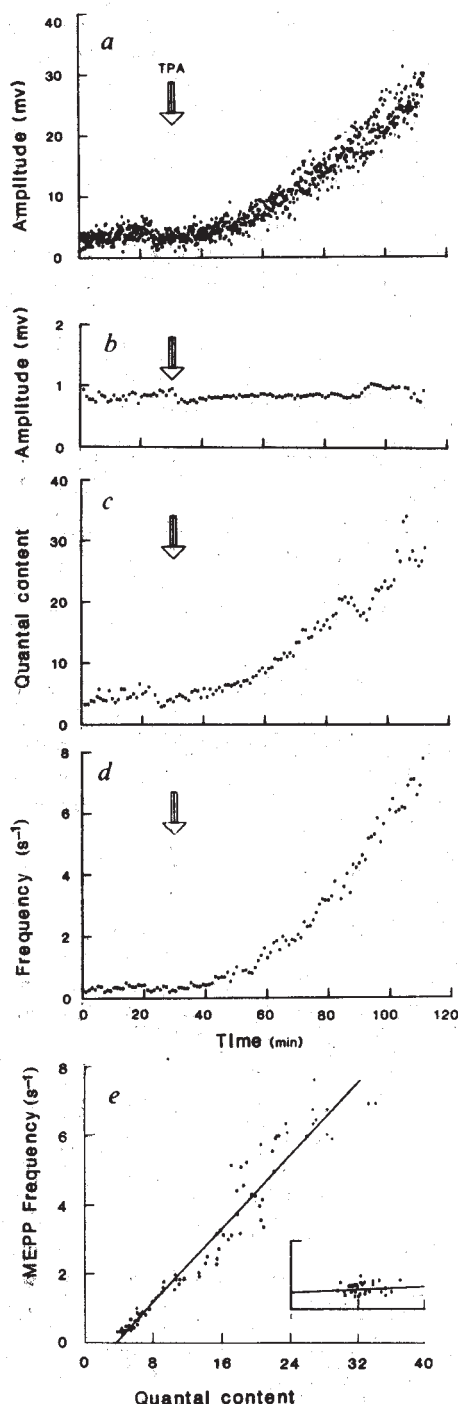


Fig. 2 The presynaptic nature of TPA action. *a*, e.p.p. amplitudes as a function of time in control Ringer solution and in solution containing TPA (162 nM). The test solution was applied at a time indicated by the arrow. Each point represents an average amplitude of three successive e.p.p.s. (The shapes of the e.p.p.s in this particular experiment are given in Fig. 1.) No corrections were made for non-linear summation. *b*, m.e.p.p. amplitudes in control Ringer solution (0–30 min) and in TPA solution (30 min onwards). Each point represents the average amplitude of m.e.p.p.s measured during one-minute intervals. *c*, Quantal content (q.c.) as a function of time. Each point represents the average q.c. per one-minute bin (20 successive e.p.p.s). Quantal contents were calculated by dividing the average e.p.p. amplitudes by the average m.e.p.p. amplitudes at the same one-minute time periods. *d*, m.e.p.p. frequency measured before and after applying TPA. Each point represents the average frequency during a one-minute interval. *e*, Correlation between m.e.p.p. frequency and q.c. during perfusion with TPA solution: $r = 0.97$. Inset demonstrates the lack of correlation between the same parameters in control solution: $r = 0.13$.

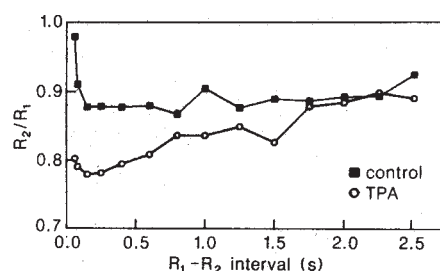


Fig. 3 Synaptic depression before (■) and after (○) the addition of TPA. The *in vitro* cutaneous pectoris nerve-muscle preparation was perfused, at room temperature, with Ringer solution containing: 115 mM NaCl, 2 mM KCl, 3.6 mM CaCl₂, 5 mM HEPES buffer (pH 7.3). Muscle contraction was abolished by pretreatment with 0.75 $\mu\text{g ml}^{-1}$ α -Bungarotoxin for 15 min (this reduces post synaptic sensitivity). The synaptic potentials, recorded by conventional electrophysiological techniques, were collected by a PDP 11-23 MINC computer. The pattern of nerve stimulation was composed of pairs of stimuli followed by a constant rest period of 20 s, allowing full recovery from depression between pairs. The interval between the conditioning pulse (R_1) and the test pulse (R_2), was varied randomly by the computer. To obtain an average depression curve, the sequence of intervals was repeated 25 times (2 h of recording for each curve). The depression curve in TPA was obtained three hours after applying the drug. Similar results were found in preparations blocked by 1.8 $\mu\text{g ml}^{-1}$ d-tubocurarine and with 1.8 mM CaCl₂ in the Ringer solution.

increase occurred in the quantal content (q.c.) of the evoked potentials (Fig. 2c). The enhancement in m.e.p.p. frequency (Fig. 2d) followed the same time course as the growth in e.p.p. amplitude, with the onset of both occurring about 20 minutes after applying TPA. Note the very good correlation between the increase in the quantal content and the increase in MEPP frequency (Fig. 2e). The finding that e.p.p. amplitude and m.e.p.p. frequency increase together with only a small change in m.e.p.p. amplitude, strongly supports the view that the action of TPA is presynaptic²⁵.

A low-activity derivative of TPA, 4-O-M-TPA^{26,27} was used as a control for both the active agent (TPA) and the solvent (DMSO). It did not affect evoked or spontaneous release at a concentration of 158 nM (100 ng ml⁻¹), in a solution containing 0.01% DMSO, when applied for up to 240 minutes (six experiments).

The results in Figs 1 and 2 clearly show that the nerve terminal increases secretion on PKC activation, as do other secreting systems. Changes in spontaneous release of transmitter due to PKC activation at the neuromuscular junction have been reported^{28,29}. The effects presented here show that activation of PKC affects nerve-stimulation-induced synaptic potentials. The results presented in Figs 1 and 2 also imply that the mode of action of PKC activation is presynaptic. The effect on evoked transmitter release could in principle be explained by an increase in calcium current^{14,15}, or the inhibition of the potassium conductance^{15,30,31}, as observed to occur with phorbol esters. However, due to the parallel increases in quantal content and m.e.p.p. frequency (Fig. 2c,d,e), we do not think that these are the major results of PKC activation.

The frequency of nerve activation is an important modulating parameter in the signalling of the nervous system. Because of the dramatic effect of TPA on evoked transmitter release, we examined the action of TPA on synaptic depression, one of the frequency modulation phenomena. In synaptic depression, the second of a pair of stimuli delivered to the motor nerve releases less transmitter than the first pulse³²⁻³⁶. This reduction in the probability of transmitter release, which occurs at high quantal contents, develops rapidly and reaches a minimum after 200–500 milliseconds; thereafter, it recovers gradually to the prestimulus level. Figure 3 shows that TPA significantly enhances depression.

This deeper synaptic depression may result from the observed increase in quantal content (Fig. 2) or may reflect a change in the release mechanism induced by TPA.

Activation of protein kinase C has profound effects on synaptic transmission. This is seen not only during direct activation of the synapse by a single stimulus, but also in more subtle modulatory phenomena such as synaptic depression. We do not

know whether the activation of kinase C affects presynaptic conductances or the secretory process itself. We speculate that the observed effects are linked to protein phosphorylation, the importance of which in neural activity is well recognized³⁷.

Supported in part by the Muscular Dystrophy Association, the Israeli Academy of Sciences and Humanities, the Council for Tobacco Research and the Bay Foundation.

Received 25 April; accepted 7 October 1986.

- Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315–321 (1984).
- Hirasawa, K. & Nishizuka, Y. A. *Rev. Pharmac. Toxicol.* **25**, 147–170 (1985).
- Katz, B. *Sherrington Lecture No. 10* (Liverpool University Press, 1969).
- Katz, B. & Miledi, R. *J. Physiol., Lond.* **203**, 689–706 (1969).
- Augustine, G. J., Charlton, M. P. & Smith, S. J. *J. Physiol., Lond.* **367**, 163–181 (1985).
- Rahamimoff, R., Lev-Tov, A., Meiri, H., Rahamimoff, H. & Nussinovitch, I. *Monogr. Neural Sci.* **7**, 3–18 (1980).
- Castagna, M. *et al. J. biol. Chem.* **257**, 7847–7851 (1982).
- Rink, T. J., Sanchez, A. & Hallan, T. J. *Nature* **305**, 317–319 (1983).
- Kojima, I., Lippes, H., Kojima, K. & Rasmussen, H. *Biochem. biophys. Res. Commun.* **116**, 552–562 (1983).
- Zawalich, W., Brown, C. & Rasmussen, H. *Biochem. biophys. Res. Commun.* **117**, 448–455 (1983).
- Tanaka, C., Taniyama, M. & Kusunoki, M. *FEBS Lett.* **175**, 165–169 (1984).
- Katakami, Y., Kaibuchi, K., Sawamura, M., Takai, Y. & Nishizuka, Y. *Biochem. biophys. Res. Commun.* **121**, 573–578 (1984).
- Valencia, M. M., Conte, D. & Negro-Vilar, A. *Brain Res. Bulletin* **15**, 657–659 (1985).
- DeRiemer, S. A., Strong, J. A., Albert, K. A., Greengard, P. & Kaczmarek, L. K. *Nature* **313**, 313–316 (1985).
- Farley, J. & Auerbach, S. *Nature* **319**, 220–223 (1986).
- Rane, S. G. & Dunlap, K. *Proc. natn. Acad. Sci. U.S.A.* **83**, 184–188 (1986).

- Kraft, A. S. & Anderson, W. B. *Nature* **301**, 621–623 (1983).
- Wolf, M., LeVine, H., May, W. S., Cuatrecasas, P. & Sahyoun, N. *Nature* **317**, 546–549 (1985).
- Sano, K., Takai, Y., Yamanishi, J. & Nishizuka, Y. *J. biol. Chem.* **258**, 2010–2013 (1983).
- Kelleher, D. J., Pessin, J. E., Ruoho, A. E. & Johnson, G. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4316–4320 (1984).
- Leeb-Lundberg, L. M. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5651–5655 (1985).
- Burgess, S. K. *et al. J. Cell Biol.* **102**, 312–319 (1986).
- DeRiemer, S. A., Greengard, P. & Kaczmarek, L. K. *J. Neurosci.* **5**, 2672–2676 (1985).
- Geron, N. & Meiri, H. *Biochim. biophys. Acta* **819**, 258–262 (1985).
- del Castillo, J. & Katz, B. *J. Physiol., Lond.* **124**, 560–573 (1954).
- Cuthbertson, K. S. R. & Cobbold, P. H. *Eur. J. cell Biol.* **14**, Suppl. 1 (1983).
- Cuthbertson, K. S. R. & Cobbold, P. H. *Nature* **316**, 541–542 (1985).
- Publicover, S. J., *Brain Res.* **333**, 185–187 (1985).
- Eusebi, F., Molinaro, M. & Caratsch, C. G. *Pflügers Arch. ges. Physiol.* **406**, 181–183 (1986).
- Baraban, J., Snyder, S. & Alger, B. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2538–2542 (1985).
- Alkon, D. L. *et al. Biochem. biophys. Res. Commun.* **134**, 1245–1253 (1986).
- Eccles, J. C., Katz, B. & Kuffler, S. W. *J. Neurophysiol.* **4**, 362–387 (1941).
- Takeuchi, A. *Jap. J. Physiol.* **8**, 102–113 (1958).
- Otsuka, M., Endo, M. & Nonomura, Y. *Jap. J. Physiol.* **12**, 573–584 (1962).
- Thies, R. E. *J. Neurophysiol.* **28**, 427–442 (1965).
- Betz, W. J. *J. Physiol. Lond.* **206**, 629–644 (1970).
- Nestler, E. J., Walaas, S. I. & Greengard, P. *Science* **225**, 1357–1364 (1984).

Activation of a calcium-dependent photoprotein by chemical signalling through gap junctions

K. Dunlap, K. Takeda* & P. Brehm

Tufts University School of Medicine, Department of Physiology, Boston, Massachusetts 02111, USA

* Université Louis Pasteur, Laboratoire de Pharmacodynamie, Strasbourg 67048, France

In the hydrozoan coelenterate *Obelia geniculata*, epithelial cell action potentials trigger light emission from photocyte effector cells^{1,2} containing obelin, an endogenous calcium-activated photoprotein^{3–5}. As this luminescence is blocked by the removal of extracellular calcium it seemed likely that calcium entry via voltage-gated channels in the photocyte membrane would account for the light emission. However, no inward calcium current was detected in whole cell recordings from dissociated photocytes and depolarization of isolated photocytes produced no luminescence. In contrast, a voltage-dependent calcium current was recorded from non-luminescent support cells, and activation of this current triggered luminescence in an adjacent photocyte. Surprisingly, light emission was abolished when the gap junctions between the photocyte and support cell were blocked. We conclude that calcium entry into support cells leads to light emission from neighbouring photocytes via chemical signalling through intercellular gap junctions.

In an intact colony of *Obelia*, an epithelial action potential can be elicited by mechanical, chemical, or electrical stimulation¹. This slowly propagating spike coordinates a range of behavioural responses from the animal, one of which is light emission from photocytes^{1,2,6}. The light is produced by a calcium-activated photoprotein, obelin^{3–5}, located in the photocyte⁶. In *Obelia*, photocytes are only a small fraction of the total cell number and, except in feeding polyps, are scattered throughout the endodermal tissue of the colony (see cover). Since the photocytes are inaccessible *in situ*, it was necessary to dissociate the tissue mechanically to study the mechanisms underlying the electrical control of luminescence. Following dissociation, photocytes were identified among the other cells, which they resembled morphologically, by the green fluorescence^{5,7} associated with obelin^{2,6}.

Since the ionic basis of the epithelial action potential is not known, we first examined the voltage-dependent channels in non-luminescent support cells by the whole cell recording technique. Action potentials can be elicited from support cells under current clamp in artificial sea water (ASW) containing either calcium, barium (Fig. 1a), or strontium. Voltage clamp analysis of these cells, isolated or in small clusters, revealed a slowly inactivating inward current after step depolarizations beyond –30 mV (Fig. 1a). The inward current was independent of external Na, absent in Ca- or Ba-free ASW, and blocked by 1 mM external Cd, indicating that activation of a voltage-dependent Ca channel⁸ underlies the action potential in support cells.

When recordings were made from either isolated photocytes or photocytes in small cell clusters, no depolarization-activated inward current was observed in either Ca- or Ba-containing ASW ($n = 59$). At voltages more positive than –20 mV, a transient outward current was elicited (Fig. 1b). This current resembled A-type K currents⁹ seen in other systems: it was reversibly blocked by 1 mM 4-aminopyridine, inactivated by conditioning depolarizations and unaffected by substitution of Ba for external Ca. Even when the outward current was suppressed, no inward current was measured in photocytes. That depolarization of isolated photocytes, either electrically ($n = 7$; Fig. 2a) or with 360 mM KCl-containing ASW ($n = 29$; Fig. 2c) does not result in light emission is consistent with the absence of a voltage-dependent Ca current in photocytes. However, when the Ca-ionophore A23187 was added to the bath, light emission from photocytes was observed as long as the extracellular solution contained Ca, showing that isolated photocytes still contain viable photoprotein.

By contrast, luminescence of photocytes which were coupled to one or more support cells (Fig. 2d) was observed following high-K stimulation of cell clusters. Several observations indicate that Ca entry in the support cell is responsible for activation of the photoprotein in the neighbouring photocytes (1) Depolarization of the support cell to levels sufficient to activate inward current resulted in light emission from the photocyte in 15 of 18 clusters tested (Fig. 2b). (2) Increasing the magnitude of the depolarization to very positive potentials, thereby reducing Ca current as the Ca equilibrium potential is approached, leads to a progressive reduction in the light response. (3) Concentrations of Cd sufficient to block inward Ca current (Fig. 3a top) inhibit light production after high K depolarization of cell clusters

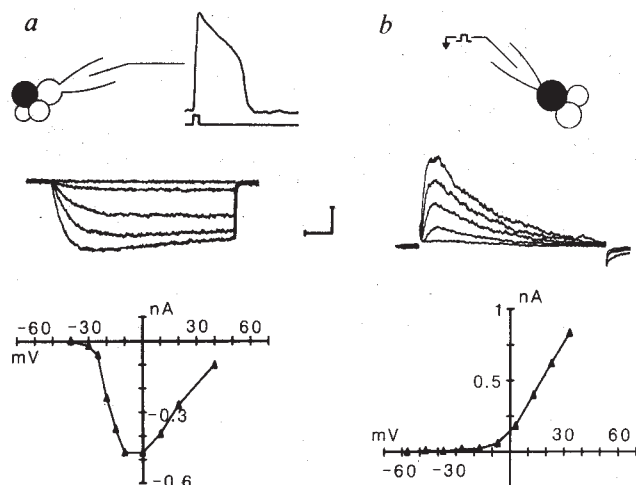


Fig. 1 Membrane currents in *Obelia* cells. Top panel, schematic representation of the configuration in recording from non-luminescent support cells (a) or photocytes (b), with the photocyte shaded to indicate its position relative to the electrode. In current clamp, action potentials (a) can be recorded from the non-luminescent cells. The 500 ms action potential illustrated here was recorded in 40 mM Ba ASW. Middle panel, leak-subtracted currents from a support cell (a) for step depolarizations of 30, 35, 40, 45 and 50 mV from a holding potential of -60 mV and from a photocyte (b) for depolarizations of 60, 70, 80, 90 and 100 mV from a holding potential of -67 mV. Calibrations: 200 pA, 30 ms. Bottom panel, complete I-V relations for the support cell (a) and photocyte (b). **Methods.** Animals were dissociated mechanically by gentle trituration through wide bore (4 mm) pipettes. Photocytes were identified using FITC fluorescence optics following illumination with a Hg source. The 490 nm excitation light intensity was attenuated by a neutral density filter and exposure time was minimized to avoid cell damage. Recordings illustrated were made at room temperature in an artificial sea water containing: 377 mM NaCl, 9 mM KCl, 40 mM BaCl₂, 40 mM MgCl₂, 2 mM NaHCO₃, pH 7.7, with an internal pipette solution composed of: 140 mM KCl, 696 mM glucose, 1 mM CaCl₂, 11 mM EGTA, 10 mM HEPES, pH 7.0 using standard patch clamp techniques.

(Fig. 3a, below). (4) Removal of external Ca blocks K-induced luminescence.

A possible role of gap junctions in Ca signalling between support cells and photocytes is suggested by the observation that photocytes are electrically coupled to their neighbouring support cells (Fig. 3b), although the extracellular stimulation technique used in this experiment did not allow a quantitative study of gap junctional conductance. Treatment of cell clusters with agents known to block gap junctions in a variety of preparations^{10,11} inhibits luminescence in *Obelia*. Low pH sea water (sodium acetate buffer, pH 6.5) and intermediate-length alcohols (250 μ M octanol or 500 μ M heptanol) reversibly and rapidly blocked K-stimulated light emission (Fig. 4).

Alcohol treatment of clusters of *Obelia* support cells effectively uncouples gap junctions between cells. Lucifer yellow dye injected into one cell of a cluster spread readily between cells in 11 of 13 untreated clusters, whereas no dye spread was observed in 7 clusters which were treated with octanol. Whole-cell recording from support cells showed that octanol did not block Ca current, so that the luminescence block was not a consequence of decreased Ca entry. Octanol-treated photocytes can still luminesce when permeabilized with ionophore A23187, indicating that octanol has no deleterious effects on the photoprotein itself. Although dye coupling experiments cannot be used to investigate the junctions between photocytes and their immediate neighbours directly, because the fluorescence emission of obelin interferes with that of Lucifer yellow, it seems unlikely that octanol blocks only junctions between support cells.

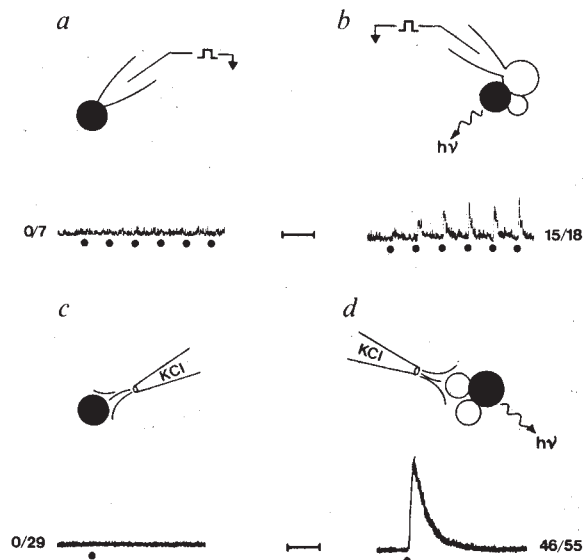


Fig. 2 Luminescence response of photocytes, either isolated (a and c) or in small clusters of cells (b and d). The method of stimulation is represented schematically at the top of each panel, with the photocyte shaded; the lower trace represents the light emission from photocytes following electrical (a and b) or high KCl-ASW (c and d) depolarization. Dots, times of stimulation; numbers to the side of each trace, proportion of cells that produced light. Calibrations, 400 ms (a and b), 2 s (c and d); emitted light was uncalibrated (same gain, a-d). **Methods.** Photocyte light emission was measured with an EMI end window photomultiplier tube attached to the camera port of a Zeiss IM35 inverted microscope. The standard ASW contained 10 mM CaCl₂. The cells in a and b were depolarized to 0 mV under whole cell clamp (Fig. 1, legend). The high-KCl ASW, with 360 mM KCl in place of NaCl, was applied locally with a pressurized puffer pipette (c and d).

We conclude that chemical signalling through gap junctions between photocytes and support cells mediates the luminescence response in *Obelia*. Because the photoprotein is activated directly by Ca, the simplest interpretation of the data is that the signal passing through the gap junction is Ca itself. In several preparations, elevated internal Ca has been shown to block rather than permeate gap junctions¹², but the internal concentration required for the block is not known. Our data do not exclude such a blocking action at high internal Ca but suggest that Ca actually moves through the gap junctions at lower concentrations.

The kinetics of light emission are consistent with Ca diffusion between cells controlling photoprotein activation. The 10–80 ms delay between the stimulation of calcium current in support cells and the onset of luminescence from the neighbouring photocyte probably reflects time for calcium to diffuse between cells rather than for activation of the photoprotein since obelin responds to calcium within 1 ms *in vitro*. The diffusion time for Ca within photocytes may not be significant since the photoprotein is located throughout the cytoplasm and so is in close proximity to Ca entering through the gap junction. The fact that only localized regions of the photocyte luminesce in response to electrical stimulation of neighbouring cells (unpublished observations by P. Brehm and S. Smith) is consistent with limited Ca diffusion in the photocyte.

Although our results are consistent with the hypothesis that Ca is the intracellular messenger mediating luminescence, involvement of a Ca-dependent transcellular signal which secondarily releases internal Ca from the photocytes cannot be excluded. However, since obelin appears to be a cytoplasmic, non-compartmentalized protein, the absence of ionophore-induced luminescence from photocytes bathed in Ca-free extracellular solutions argues against a second messenger

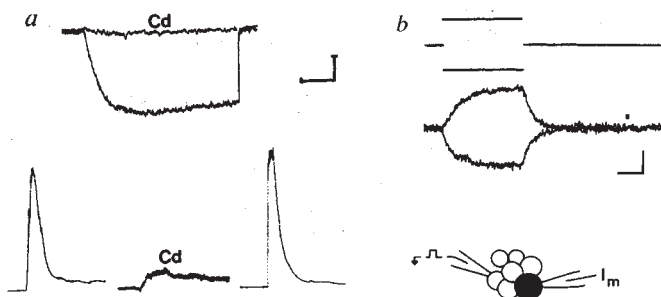


Fig. 3 *a*, Cadmium block of inward current and luminescence. Above, support cell inward current in 40 mM Ba-ASW before (lower trace) and during puff application of 1 mM CdCl₂ in 40 mM Ba ASW, evoked by step depolarizations to -15 mV from a holding potential of -60 mV under whole cell clamp. Calibrations: 100 pA, 35 ms. Below, 450 mM KCl-stimulated light production from suspensions of cell clusters in standard ASW before Cd application (left trace), during 1 mM Cd (middle) and after extensive washing out of Cd (right). Time calibration: 2 s; light emission uncalibrated. Gain of the middle trace is $\times 10$ relative to others. *b*, Electrical coupling between photocytes and neighbouring support cells. Whole cell recording of current (I_m) from a photocyte is shown in response to extracellular stimulation of the cluster of surrounding support cells. Both polarities of stimulation result in measurable current flow in the photocyte. Recording configuration illustrated schematically in the lower panel, with the photocyte shaded. Calibration: 1 nA, 1 V and 10 ms.

mediated release of internal Ca from the photocyte. Regardless of the exact mechanism underlying the rise in internal photocyte Ca concentration, we have shown that gap junctions in excitable tissue pass functionally important chemical signals. Although such a role for intercellular channels has been previously suggested¹³, our results provide experimental support, in a natural biological system, for gap junction chemical signalling in intercellular communication. This might represent a third basic mechanism resulting in an increase in intracellular Ca activity: in addition to inward Ca current through membrane channels⁸

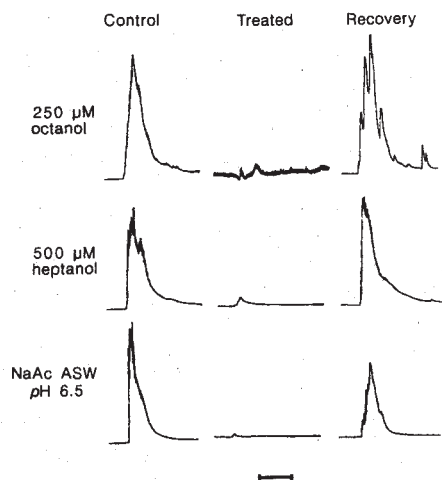


Fig. 4 Effect of gap junction uncouplers on light emission from suspensions of cell clusters stimulated by 450 mM KCl. The controls (left column) represent clusters before treatment. The treated clusters (middle) were incubated with the uncouplers and tested within 2 min. The alcohols were added to standard ASW and low-pH ASW was made by substituting Na acetate for NaCl and adjusting pH to 6.5 with HCl. Recovery from the uncoupling treatment (right) was tested following extensive washing of treated clusters with ASW. Light emission was measured as in Fig. 3, legend. Horizontal calibration: 2 s; emitted light uncalibrated. The gain was constant throughout except for octanol treated cells (middle trace) $\times 10$.

or Ca release from intracellular stores¹⁴, we find that Ca may enter an effector cell via neighbouring support cells.

We thank P. A. V. Anderson, L. A. Jaffe and R. T. Kado for critical comments and discussion and G. Mandel and D. Youngman for editorial help. This work was done at the Marine Biological Laboratory, Woods Hole, Massachusetts, USA and was supported by NSF grant 8503159. Additional support to K.T. was provided by the CNRS (UA600) and MBL Summer Fellowships.

Received 24 July; accepted 23 October 1986.

1. Morin, J. G. & Cooke, I. *J. exp. Biol.* **54**, 707-721 (1971).
2. Morin, J. G. in *Coelenterate Biology: Reviews and New Perspectives*, 397-438 (Academic, New York, 1974).
3. Campbell, A. K., *Biochem. J.* **143**, 414-418 (1974).
4. Stephenson, D. G. & Sutherland, P. J. *Biochim. biophys. Acta* **678**, 65-75 (1981).
5. Morin, J. G. & Hastings, J. W. *J. cell. Physiol.* **77**, 313-318 (1971).
6. Morin, J. G. & Reynolds, G. *Biol. Bull.* **147**, 397-410 (1974).
7. Johnson, F. et al. *J. cell. comp. Physiol.* **60**, 85-104 (1963).
8. Hagiwara, S. & Byerly, L. A. *Rev. Neurosci.* **4**, 69-125 (1981).
9. Rogawski, M. A. *Trends Neurosci.* **8**, 214-219 (1985).
10. Johnston, M. F., Simon, S. A. & Ramon, F. *Nature* **286**, 498-500 (1980).
11. Spray, D. C., White, R. L., Campos de Carvalho, A., Harris, A. L. & Bennett, M. V. L. *Biophys. J.* **45**, 219-230 (1984).
12. Spray, D. C., Stern, J. H., Harris, A. L. & Bennett, M. V. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 441-445 (1982).
13. Lawrence, T. S., Beers, W. H. & Gilula, N. B. *Nature* **272**, 501-506 (1978).
14. Rasmussen, H. & Goodman, D. B. P. *Physiol. Rev.* **57**, 421-509 (1977).

A new mechanism for induced vitamin D deficiency in calcium deprivation

M. R. Clements*, L. Johnson & D. R. Fraser

Dunn Nutritional Laboratory, University of Cambridge and Medical Research Council, Milton Road, Cambridge CB4 1XJ, UK

Synthesis of vitamin D in the skin in response to ultraviolet light is the main determinant of vitamin D status in man¹ and it is therefore surprising that rickets and osteomalacia, clinical signs of vitamin D deficiency, remain common in tropical and subtropical countries². Skin pigmentation can reduce vitamin D formation³ but this is a negligible limitation in people exposed to abundant ultraviolet light^{4,5}. Earlier studies in animals⁶ and man⁷ suggested that another environmental factor, the low calcium/high cereal diet typical of susceptible populations², might affect the efficiency of vitamin D utilization. We show here in rats that the rate of inactivation of vitamin D in the liver is increased by calcium deprivation. The effect is mediated by 1,25-dihydroxyvitamin D, produced in response to secondary hyperparathyroidism^{8,9}, which promotes hepatic conversion of vitamin D to polar inactivation products that are excreted in bile. This finding has widespread implications both for understanding the pathogenesis of endemic rickets and in that it provides a unifying mechanism for the development of vitamin D deficiency in many clinical disorders.

Plasma levels of tracer radio-labelled 25-hydroxyvitamin D (25(OH)D) exhibit a biphasic exponential decline¹⁰ after intravenous injection. In rats the initial rapid phase is followed by a longer phase with a half-time ($t_{1/2}$ -25(OH)D) of several days. When vitamin D supply is limited, the long-phase $t_{1/2}$ -25(OH)D indicates the rate of utilization of vitamin D. We measured $t_{1/2}$ -25(OH)D under conditions of varying calcium supply and found a linear relationship between $t_{1/2}$ -25(OH)D and log dietary calcium concentration (Fig. 1a). With deprivation of dietary calcium, mean $t_{1/2}$ -25(OH)D decreased from 15.4 days (0.64% Ca) to 9.2 days (0.02% Ca) (Fig. 1b). A similar decrease ($t_{1/2} = 10.3 \pm 0.6$ days, mean $\pm$ s.e.m., $P < 0.001$, $n = 4$)

* Present address: University Department of Medicine, Royal Infirmary, Manchester M13 9WL, UK.

Fig. 1 Levels of 25-hydroxyvitamin D in rats. Male Norwegian hooded rats, housed in the absence of ultra-violet light, were fed on vitamin D-deficient diets³⁵ containing 0.02%–0.64% calcium. Adequate and defined vitamin D status was obtained by subcutaneous injection of 5–10 μ g cholecalciferol (D_3) in propylene glycol and 10 days later the 25(OH)D pool was labelled by intracardial injection of 0.3 μ Ci [$26,27\text{-}^3\text{H}$]25(OH) D_3 (20.6 Ci mmol^{-1}) in vitamin D-deficient rat serum. The elimination half-time of ^3H -25(OH) D_3 [$t_{1/2}\text{-}^3\text{H}$ -25(OH) D_3] was measured from the fourth day after its injection. Blood was collected by cardiac puncture. A single subcutaneous dose of 5 μ g D_3 was effective in maintaining adequate vitamin D status over a 4 week period. Varying the dose in the range 5–125 μ g had no significant effect on $t_{1/2}\text{-}^3\text{H}$ -25(OH) D_3 and there was no change in the specific activity of ^3H -25(OH) D_3 in plasma during these measurements. **a**, Effect of dietary [Ca] g per kg of diet on $t_{1/2}\text{-}^3\text{H}$ -25(OH) D_3 ($n=3$ per group). **b**, Decline in plasma [^3H -25(OH) D_3] on 0.64% (○) and 0.02% (●) Ca diets ($P<0.03$, $n=4$ per group). **c**, Cumulative biliary excretion of ^3H in bile-duct cannulated rats on 0.64% (○) and 0.02% (●) Ca diets ($P<0.03$, $n=4$ per group). **d**, Effect of dietary phosphate content and parathyroidectomy (PTX) on decline in plasma [^3H -25(OH) D_3] on 0.64% Ca diet. Values for $t_{1/2}\text{-}^3\text{H}$ -25(OH) D_3 : 0.54% P (○) = 15.3 ± 0.1 days ($n=4$); 1.20% P (●) = 11.6 ± 0.5 ($n=4$; $P<0.001$). After PTX, $t_{1/2}\text{-}^3\text{H}$ -25(OH) D_3 = 16.7 ± 0.7 days. **e**, Effect of 1,25-(OH) $_2D_3$ subcutaneous infusion (1 ng per 100 g per h) on decline in plasma [^3H -25(OH) D_3] on 0.64% Ca diet. Values for $t_{1/2}\text{-}^3\text{H}$ -25(OH) D_3 : (i) pre-infusion = 14.5 ± 0.9 days ($n=5$); (ii) during infusion = 6.3 ± 0.1 ($P<0.001$); (iii) post-infusion = 12.3 ± 1.3 days. **f**, Effect of 1,25-(OH) $_2D_3$ subcutaneous infusion (●) on cumulative biliary excretion of ^3H in bile-duct cannulated rats on 0.64% Ca diet compared with untreated rats (○) ($P=0.01$; $n=4$ per group).

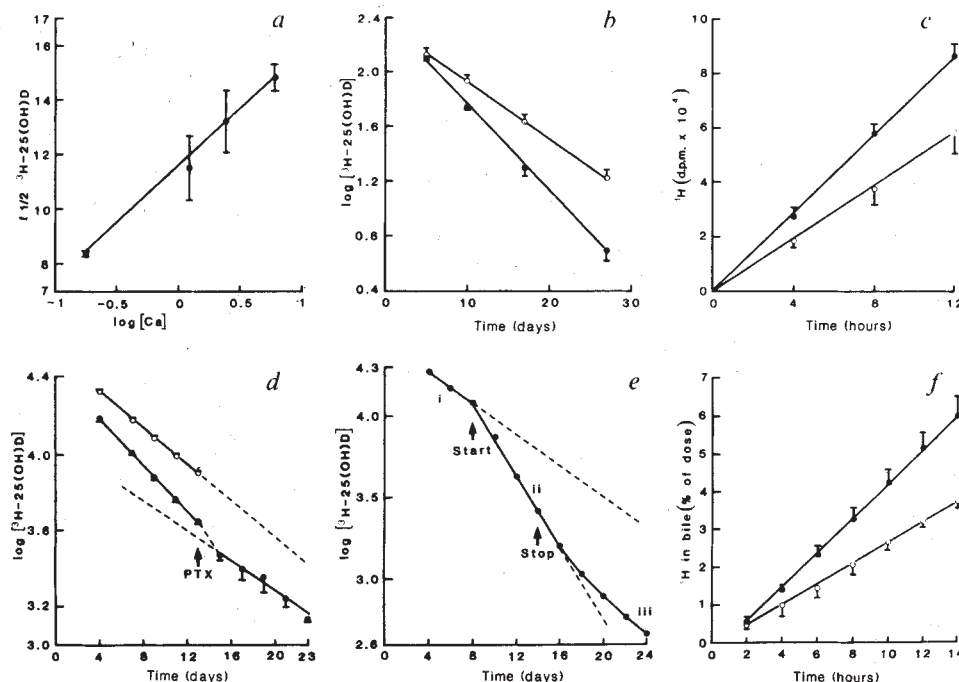


Table 1 Plasma variables in rats receiving low calcium or high phosphorus intakes

Diet and treatment	Calcium (mmol l ⁻¹)	25(OH)D (ng ml ⁻¹)	1,25(OH) $_2D_3$ (pg ml ⁻¹)	$t_{1/2}\text{-}^3\text{H}$ -25(OH) D_3 (days)	$t_{1/2}\text{-}^3\text{H}$ -1,25(OH) $_2D_3$ (hours)
0.64% Ca	2.51 $\pm$ 0.01	13.4 $\pm$ 1.4	78 $\pm$ 6 (6)	15.4 $\pm$ 0.5 (23)	34.6 $\pm$ 1.2 (6)
0.02% Ca	2.48 $\pm$ 0.07	7.9 $\pm$ 1.0**	389 $\pm$ 45*** (6)	9.2 $\pm$ 0.6*** (13)	35.0 $\pm$ 0.6 (6)
0.64% Ca + 1,25(OH) $_2D_3$ infusion	3.85 $\pm$ 0.11***	2.3 $\pm$ 0.2***	545 $\pm$ 57*** (4)	6.3 $\pm$ 0.1*** (5)	—
1.2% P	2.52 $\pm$ 0.01	6.5 $\pm$ 0.5***	101 $\pm$ 6* (6)	11.6 $\pm$ 0.5*** (4)	—
1.2% P + PTX	1.38 $\pm$ 0.04***	3.1 $\pm$ 0.4***	3 $\pm$ 1*** (7)	16.7 $\pm$ 0.7	—
0.64% Ca + phenobarbitone	—	—	—	12.6 $\pm$ 0.4*** (5)	—
0.02% Ca + phenobarbitone	—	—	—	8.5 $\pm$ 0.5*** (5)	—

Values are means $\pm$ s.e.m. Significant differences from control values (in rats receiving a diet containing 0.64% Ca and 0.54% P), are shown as: *, $P<0.02$; **, $P<0.01$; ***, $P<0.001$. Figures in parentheses show the number of values averaged.

was found when sodium phytate (5%) was included in the 0.64% Ca diet. The accelerated loss of 25(OH)D on 0.02% Ca diet could be quantitatively accounted for by an 86% greater output of labelled vitamin D metabolites in faeces of calcium-deficient compared with control rats. This reflected a similar increase in output of polar vitamin D metabolites in the bile of calcium-deficient rats (Fig. 1c).

There are a number of possible mechanisms by which calcium deficiency might enhance the degradative metabolism of 25(OH)D. The simplest explanation is that 25(OH)D is consumed by the increased production of 1,25-dihydroxyvitamin D (1,25(OH) $_2D_3$) in response to calcium deprivation¹¹ (Table 1). However, because published data indicate that 1,25(OH) $_2D_3$ production is increased only three- to fivefold in experimental calcium deficiency¹², this change appears unlikely to account for the loss of 25(OH)D. Furthermore, as the metabolic clearance of 1,25(OH) $_2D_3$ in rats is not altered by calcium deficiency (ref. 13 and Table 1), the increased production of this metabolite is reflected directly by the fivefold increase in its concentration in plasma (Table 1). Analysis of biliary metabolites following an intravenous injection of tracer [$4\text{-}^{14}\text{C}$, 1,2- ^3H]cholecalciferol

confirmed this interpretation. When 1-hydroxylation occurs, ^3H in the 1 α -position of the vitamin D molecule is lost¹⁴. Comparison of the $^3\text{H}/^{14}\text{C}$ ratio in bile with that of the dose therefore allows calculation of the maximum quantity of biliary metabolites derived from 1,25(OH) $_2D_3$ (ref. 15). This was found to

Table 2 Effect of calcium intake on biliary output of metabolites of [$4\text{-}^{14}\text{C}$, 1,2- ^3H]cholecalciferol in rats

Dietary calcium (%)	No. of values	Metabolites derived from 1,25(OH) $_2D_3$ (% of ^{14}C dose)	Other labelled metabolites (% of ^{14}C dose)
0.64	4	0.28 $\pm$ 0.07	3.57 $\pm$ 0.39
0.02	4	1.44 $\pm$ 0.26 ($P<0.01$)	5.83 $\pm$ 0.3 ($P<0.01$)

Rats received intracardial injections of a mixture of [$4\text{-}^{14}\text{C}$]cholecalciferol (0.2 μ Ci; 57 mCi mmol^{-1}) and [1,2- ^3H]cholecalciferol (1 μ Ci; 18.3 mCi mmol^{-1}) 48 h before bile duct cannulas were surgically implanted. Values are means $\pm$ s.e.m. of ^{14}C -labelled metabolites collected over 12 h. Values of P obtained as in Table 1.

Table 3 Effect of calcium deficiency on aspects of hepatic function

Dietary calcium (%)	Liver blood flow (% cardiac output)		Short-term hepatic uptake of ^3H -25(OH) D_3 , 30 min after injection intracardially (% of dose)	Hepatic cytochrome P-450 (mmol per g body weight)	Glucuronides in bile (% of ^3H -labelled metabolites)
	Hepatic artery	Portal vein			
0.64	2.28 ± 0.14	13.49 ± 0.47 (5)	2.95 ± 0.21 (3)	0.249 ± 0.03 (5)	16.1 ± 2.2 (4)
0.02	1.57 ± 0.16	12.2 ± 0.82 (5)	3.01 ± 0.46 (3)	0.246 ± 0.017 (5)	17.0 ± 2.9 (4)
0.64 +phenobarbitone	—	—	3.57 ± 0.33 (3)	1.087 ± 0.069 (5) ($P < 0.001$)	17.7 ± 1.1 (4)

Figures in parentheses indicate the number of values averaged. Values of P obtained as in Table 1.

be fivefold higher in bile from calcium-deficient rats. However, the biliary output of ^{14}C -labelled metabolites other than those derived from $1,25(\text{OH})_2\text{D}$ was still 63% greater than controls (Table 2). Thus increased production of $1,25(\text{OH})_2\text{D}$ accounted for only a small part of the observed disappearance of $25(\text{OH})\text{D}$.

There was no change in the serum calcium of rats on the low calcium diet (Table 1) so it seemed unlikely that extracellular calcium could be affecting directly the enzymes involved in the hepatic degradation of $25(\text{OH})\text{D}$. The possibility remained that the elevated levels of either parathyroid hormone (PTH) or $1,25(\text{OH})_2\text{D}$ might alter the hepatic metabolism of $25(\text{OH})\text{D}$ in calcium-deprived animals.

To investigate this we induced moderate hyperparathyroidism in rats by feeding them a high phosphate diet^{16,17} (1.2% P) which decreased mean $t_{1/2}$ - $25(\text{OH})\text{D}$ from 15.3 to 11.6 days (Table 1). After parathyroidectomy, $t_{1/2}$ - $25(\text{OH})\text{D}$ was restored to control values, implicating PTH in the enhanced degradation of $25(\text{OH})\text{D}$ (Fig. 1d). PTH increases hepatic blood flow in dogs¹⁸ and so might increase clearance of $25(\text{OH})\text{D}$ in the liver directly, but in rats there was no effect of calcium deficiency on the proportion of cardiac output flowing through the hepatic artery or portal vein (estimated using radio-labelled microspheres¹⁹), or on the short-term (30 min) hepatic uptake of ^3H - $25(\text{OH})\text{D}$ (Table 3). A similar study in man showed no change in liver blood flow measured by indocyanine green clearance²⁰ after parathyroidectomy for hyperparathyroidism (unpublished data).

The alternative hypothesis, that the effect of hyperparathyroidism on the hepatic metabolism of $25(\text{OH})\text{D}$ was mediated indirectly by $1,25(\text{OH})_2\text{D}$, was investigated by infusing rats with $1,25(\text{OH})_2\text{D}$ by means of subcutaneous osmotic minipumps (Alzet). This produced plasma levels of $1,25(\text{OH})_2\text{D}$ comparable to those found on 0.02% Ca diet (Table 1) and caused a rapid decline in mean $t_{1/2}$ - $25(\text{OH})\text{D}$ from 14.5 to 6.3 days, reverting to 12.3 days on termination of the infusion (Fig. 1e). This treatment also caused a 60% increase in biliary excretion of $25(\text{OH})\text{D}$ degradation metabolites compared with controls (Fig. 1f). We conclude that elevation of PTH secretion stimulates increased production of $1,25(\text{OH})_2\text{D}$ which acts directly or indirectly to enhance inactivation of $25(\text{OH})\text{D}$ in the liver. Whether or not PTH itself has an independent effect on the hepatic metabolism of $25(\text{OH})\text{D}$ remains uncertain.

The specific mechanism of action of $1,25(\text{OH})_2\text{D}$ on the hepatic metabolism of $25(\text{OH})\text{D}$ is unclear because little is known about the degradative enzymes involved. Hepatic glucuronidation facilitates the biliary excretion of molecules by increasing their polarity and size, and glucuronide conjugates of vitamin D metabolites have been reported in bile²¹. However, the proportion of labelled biliary metabolites present as glucuronides was not affected by calcium deficiency (Table 3).

Treatment with the anticonvulsant phenobarbitone, which induces hepatic cytochrome P-450 enzymes, is often associated with decreased levels of $25(\text{OH})\text{D}$ in plasma²² and can lead to the development of osteomalacia. Both phenobarbitone treatment (1 mg ml^{-1} in drinking water) and calcium deprivation

separately lowered $t_{1/2}$ - $25(\text{OH})\text{D}$ in rats (Table 1) and increased the output of polar vitamin D metabolites in bile. Moreover, both treatments together had an additive effect, suggesting that they operated through different mechanisms. In support of this, the hepatic microsomal cytochrome P-450 content²³ was not affected by calcium deficiency whereas there was a fourfold increase in the hepatic cytochrome P-450 levels in rats supplied with phenobarbitone (Table 3). Therefore it appears likely that the effect of excess $1,25(\text{OH})_2\text{D}$ on the catabolism of $25(\text{OH})\text{D}$ is due to activation of existing hepatic oxidative enzymes. This may be an exaggerated response of a normal function for $1,25(\text{OH})_2\text{D}$ in liver but there is no evidence to suggest that it represents a feed-back control mechanism to regulate plasma levels of $25(\text{OH})\text{D}$. Likewise, the enhanced inactivating metabolism may not be specifically directed against $25(\text{OH})\text{D}$ but may also apply to other substrates of hepatic catabolic pathways. It is also possible that $1,25(\text{OH})_2\text{D}$ may enhance hepatic inactivation of $25(\text{OH})\text{D}$ without acting directly in the liver. However, in acute experiments (unpublished data) we have demonstrated that the rate of output of ^3H - $25(\text{OH})\text{D}$ metabolites in bile is increased within 4 h of raising the plasma concentration of $1,25(\text{OH})_2\text{D}$. Such a rapid response suggests that $1,25(\text{OH})_2\text{D}$ acts directly, but verifications of this can only come from experiments with isolated liver preparations *in vitro*.

Although the $1,25(\text{OH})_2\text{D}$ receptors found in most mammalian cells²⁴ have not yet been demonstrated in hepatocytes, reports that $1,25(\text{OH})_2\text{D}$ increases hepatocyte cytosolic calcium levels²⁵ and stimulates a protein kinase in rat liver²⁶ suggest that this organ could be a target for this vitamin D metabolite.

The finding that calcium deficiency promotes vitamin D deficiency could be important in understanding the pathogenesis and epidemiology of endemic rickets. The half-time of $25(\text{OH})\text{D}$ in man is also shortened in primary hyperparathyroidism (unpublished data) and we suggest that the effect of $1,25(\text{OH})_2\text{D}$ on the hepatic metabolism of $25(\text{OH})\text{D}$ can explain the puzzling development of vitamin D deficiency in many clinical disorders characterized by hyperparathyroidism^{27,28} or calcium malabsorption. These include gastro-intestinal disease^{29,30}, resection³¹ or bypass³², chronic liver disease^{30,33}, anticonvulsant therapy²³ and morbid obesity³⁴ as well as rickets of Asian immigrants in the UK².

We thank Dr E. B. Mawer for plasma assays of $25(\text{OH})\text{D}$ and $1,25(\text{OH})_2\text{D}$.

Received 21 May; accepted 14 October 1986.

- Haddad, J. G. & Hahn, T. J. *Nature* **244**, 515-517 (1973).
- DHSS Report on Health and Social Subjects, 19 (HMSO, London, 1980).
- Clemens, T. L., Adams, J. S., Henderson, S. L. & Holick, M. F. *Lancet* **i**, 74-76 (1982).
- Stamp, T. C. B. *Proc. Nutr. Soc.* **34**, 119-130 (1975).
- Davie, M. & Lawson, D. E. M. *Clin. Sci.* **58**, 235-242 (1980).
- Ewer, T. K. *Nature* **166**, 732-733 (1950).
- Robertson, I., Ford, J. A., McIntosh, W. B. & Dunnigan, M. G. *Br. J. Nutr.* **45**, 17-22 (1981).
- Garabedian, M., Holick, M. F., DeLuca, H. F. & Boyle, I. T. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1673-1676 (1972).
- Fraser, D. R. & Kodicek, E. *Nature new Biol.* **241**, 163-166 (1973).
- Gray, R. W., Weber, H. P., Dominguez, J. H. & Lemann, J. *J. clin. Endocr. Metab.* **39**, 1045-1056 (1974).
- Ribovich, M. L. & DeLuca, H. F. *Archs Biochem. Biophys.* **188**, 145-156 (1978).
- Fox, J. & Ross, R. *J. Endocr.* **105**, 169-173 (1985).

13. Paulson, S. K. & Kenny, A. D. *Biopharmaceutics and Drug Disposition* 6, 359-372 (1985).
14. Lawson, D. E. M., Wilson, P. W. & Kodicek, E. *Nature* 222, 171-172 (1969).
15. Bell, P. A. & Kodicek, E. *Biochem. J.* 116, 755-757 (1970).
16. Anderson, G. H. & Draper, H. H. *J. Nutr.* 102, 1123-1132 (1972).
17. Siu, G. M. & Draper, H. H. *Calc. Tiss. Int.* 33, 667-672 (1981).
18. Charbonne, G. A. & Hulstaert, P. F. *Endocrinology* 95, 621-626 (1974).
19. Groszmann, R. G., Vorobioff, J. & Riley, E. *Am. J. Physiol.* 242, G156-G160 (1982).
20. Ketterer, S. G., Wiegand, B. D. & Rapaport, E. *Am. J. Physiol.* 199, 481-484 (1960).
21. Bell, P. A. & Kodicek, E. *Biochem. J.* 115, 663-669 (1969).
22. Hahn, T. J., Hendin, B. A., Scharp, C. R. & Haddad, J. G. *New Engl. J. Med.* 287, 900-904 (1972).
23. Omura, T. & Sato, R. *J. biol. Chem.* 239, 2370-2378 (1964).
24. Braidman, I. P. & Anderson, D. C. *Clin. Endocr.* 23, 445-460 (1985).
25. Baran, D. T. & Milne, M. L. *J. clin. Invest.* 77, 1622-1626 (1986).
26. Sikorska, M. & Whitfield, J. F. *Biochem. biophys. Res. Commun.* 129, 766-772 (1985).
27. Woodhouse, N. J. Y., Doyle, F. H. & Joplin, G. H. *Lancet* ii, 283-287 (1971).
28. Stanbury, S. W. *J. R. Coll. Phys. Lond.* 15, 205-217 (1981).
29. Driscoll, R. H., Meredith, S. C., Sitrin, M. & Rosenberg, I. H. *Gastroenterology* 83, 1252-1258 (1982).
30. Dibble, J. B., Sheridan, P. & Losowsky, M. S. *Q. J. Med.* 209, 119-134 (1984).
31. Morgan, S. B., Hunt, G. & Patterson, C. R. *Q. J. Med.* 39, 395-410 (1970).
32. Teitelbaum, S. L., Halverson, J. D., Bates, M., Wise, L. & Haddad, J. G. *Ann. intern. Med.* 86, 289-293 (1977).
33. Dibble, J. B., Sheridan, P., Hampshire, R., Hardy, G. J. & Losowsky, M. S. *Clin. Endocr.* 15, 373-383 (1981).
34. Bell, N. H., Epstein, S., Greend, A., Shary, J., Oexmann, M. J. & Shaw, S. *J. clin. Invest.* 76, 370-373 (1985).
35. DeLuca, H. F., Guroff, E., Steenbock, H., Reiser, F. & Mannatt, M. R. *J. Nutr.* 75, 175-180 (1961).

Characterization of murine cytolytic-helper hybrid T cell clones

Wendy L. Havran & Frank W. Fitch

Committee on Immunology and Department of Pathology, The University of Chicago, Box 414, 5841 S. Maryland Avenue, Chicago, Illinois 60637, USA

L3T4, Lyt-2 and the T-cell receptor for antigen are cell-surface molecules involved in antigen specific T cell activation. We have constructed functional murine cytolytic-helper T-cell hybrid clones to study the link between expression of cell-surface molecules and specific cell function. Three of the clones express two antigen receptors, and both Lyt-2 and L3T4, normally expressed on mutually exclusive subsets of mature T lymphocytes. The pattern of lymphokines produced by the hybrid cells in response to antigen was not controlled by the specific antigen receptor; both T-cell growth factor, produced only by the helper T-cell partner, and γ -interferon, produced only by the cytolytic T-cell partner, were secreted when either antigen receptor was stimulated. However, cytolytic activity appeared to be restricted to the recognition of antigen by the T-cell-receptor of the cytolytic partner. Thus cytotoxicity appears to be rightly linked to the antigen receptor of the cytolytic parent but lymphokine release is not tightly linked.

T lymphocytes are important in both cellular and humoral immune responses. They carry out direct effector functions such as cytotoxicity and mediate regulatory functions via secreted lymphokines. Murine T-cell surface molecules which are involved in T-cell activation include the antigen receptor, Lyt-2 and L3T4. Normally, Lyt-2 and L3T4 are expressed on mutually exclusive subsets of mature murine T cells and correlate with MHC restriction and, to a considerable extent, cell function.

To study T-cell regulation and requirements for cytotoxicity, other investigators have fused tumour cells with T cells^{1,2} or inserted membrane fragments into cells of different specificities^{3,4}. However, the fusion of two phenotypically and functionally distinct normal T cells provides a better-defined system and should permit study of the linkage between receptor specificity and specific cell function. Such a system should provide insight into the regulatory processes that control T cell responses and the mechanisms involved in differentiation and proliferation of T lymphocytes.

To study the linkage between receptor specificity and cell function, we constructed murine cytolytic-helper T-cell hybrids by fusing normal murine cytolytic T-lymphocyte (CTL) and

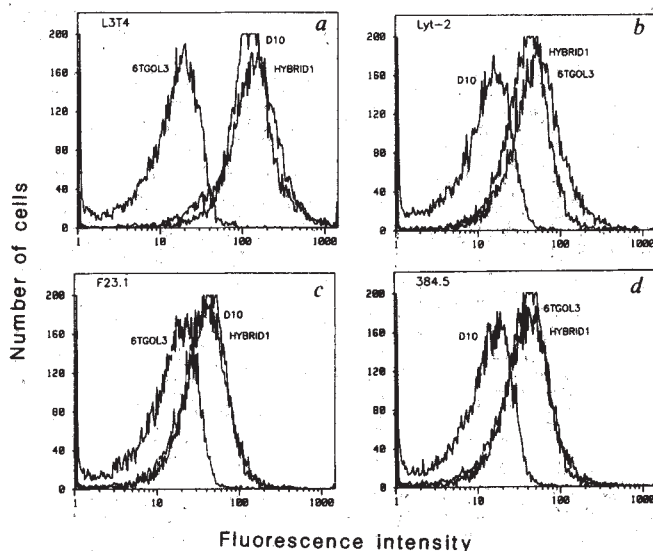


Fig. 1 Immunofluorescence analysis of hybrid 1 and the fusion partner T cell clones, 6TGOL3 and D10. Cells were centrifuged over a Ficoll-hypaque gradient, washed and distributed for fluorescence staining. 10^6 cells were incubated with saturating amounts of culture supernatant containing monoclonal antibody or control supernatant for 15 min at 4 °C. After washing, cells were incubated with *a, b*, mouse anti-rat κ chain monoclonal antibody, or *c, d*, fluorescein-conjugated goat anti-mouse Ig for an additional 15 min. Cell surface staining was quantified using a FACS analyser (B-D FACS Systems). In all experiments, propidium iodide was added to identify and exclude dead cells. Fluoresceinated second-step antibody alone stained less than 2% of the cells.

helper T-lymphocyte (HTL) clones with different antigen specificities. A variant of the murine CTL clone L3⁵ has been derived to serve as a 'universal' fusion partner. The CTL clone L3 is of C57BL/6 origin and is reactive with the L^d alloantigen⁵. Upon stimulation with concanavalin A (Con A) or antigen, L3 releases γ -interferon (IFN- γ) but not T-cell growth factor (TCGF), interleukin-3 (IL-3) or cerebrospinal fluid (CSF)^{8,9}. A HAT (hypoxanthine, aminopterin and thymidine)-sensitive, ouabain-resistant variant of this clone that retains normal L3 phenotype and specific cytotoxic activity was derived by sequentially selecting for growth in the presence of 6-thioguanine (6-TG) and 3 mM ouabain. Viable cells were cloned in the presence of 6-TG and ouabain. This clone has been designated 6TGOL3.

This 6TGOL3 clone has been successfully fused with several different HTL clones including L2 (ref. 6), HK16 (ref. 5), and D10.G4.1 (ref. 7) with similar results. Only the results of fusion with HTL clone D10.G4.1 (D10) will be discussed in detail here. Clone D10 is of AKR origin and is reactive with conalbumin in the context of I-A^k and has alloreactivity against I-A^b (ref. 7). The D10 clone does not demonstrate lytic activity in an antigen-specific or lectin-mediated cytotoxic assay¹⁰. D10 cells require IL-1 for growth and secrete TCGF but not IL-3, CSF, or IFN- γ after stimulation. The TCGF activity secreted by D10 cells appears to be due to BSF-1, rather than IL-2 (K. Bottomley and C. A. Janeway, Jr, personal communication). In these studies this lymphokine activity is measured using CTLL-2 indicator cells which also proliferate in response to BSF-1. The 6TGOL3 and D10 clones express different cell surface antigens, secrete a different array of lymphokines, have different functional activities and react with different classes of MHC antigens. 6TGOL3 and D10 cells were fused using the polyethylene glycol (PEG) method with modifications described by others for obtaining T-cell hybrids^{11,12}. Hybrid cells were isolated after 7-10 days in selective media then cloned by micromanipulation to ensure that clones originated from single cells.

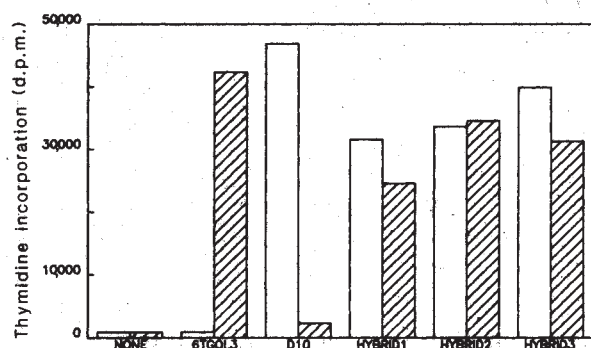


Fig. 2 Proliferation of the fusion partner T cells clones 6TGOL3 and D10 and three hybrid clones to: open bars, irradiated (1,400 rad), CA14.11.14 cells plus conalbumin; matched bars, T-cell-depleted DBA/2 spleen cells. Spleen cells were depleted of T cells by incubating with anti-Thy-1 monoclonal antibody AT83A for 30 min at 4 °C, washing, and then incubating with absorbed rabbit complement for 45 min at 37 °C. 5×10^6 cloned T cells were cultured for 36 h with 1×10^6 irradiated (1,400 rads) DBA/2 spleen cells (L^d) or CA14 ($I-A^k$ transfected) cells, previously cultured with $100 \mu\text{g ml}^{-1}$ ConA for 24 h. ^3H -thymidine (1 μCi) was added and thymidine incorporation was measured 12 h later.

Reactivity of the hybrid cells with anti-Lyt-2, anti-L3T4, and anti-T-cell receptor monoclonal antibodies was measured by immunofluorescence and is shown in Fig. 1. FACS analysis showed that most, if not all, hybrid cells expressed both Lyt-2 and L3T4. Fluorescence microscopy using fluorescein-isothiocyanate (FITC)-anti-Lyt-2 and phycoerythrin-anti-L3T4 verified that individual cells were double positive. Separate immunofluorescence staining indicated the presence of both antigen receptors on most, if not all, of the cloned hybrid cells.

We examined the ability of three HAT-resistant, ouabain-resistant hybrid clones to proliferate in response to irradiated, T cell-depleted L^d -bearing DBA/2 spleen cells (Fig. 2) or irradiated, $I-A^k$ -bearing CBA spleen cells plus conalbumin (data not shown). The 6TGOL3 parent proliferated when cultured with DBA/2 spleen cells but not when cultured with CBA spleen cells and conalbumin, while the D10 parent proliferated when cultured with CBA spleen cells and conalbumin but not when cultured with DBA/2 spleen cells. The three hybrid clones proliferated in response to both populations. The D10 clone and the hybrid clones could also proliferate when stimulated

by conalbumin presented by the $I-A^k$ transfected L cell line CA14.11.14 (CA14)¹³ (Fig. 2). Since these cells stimulated proliferation of the hybrid cells when conalbumin was included in the culture, conalbumin-pulsed CA14 cells should provide a uniform population of target cells suitable for evaluating cytolytic activity mediated through the D10 antigen receptor.

The three hybrid clones were analysed for specific lysis of L^d -bearing P815 target cells or the $I-A^k$ -transfected line CA14 after culture for 24 h in the presence of conalbumin (CA14/ca). As shown in Fig. 3, all three hybrid clones lysed P815 target cells, as did the 6TGOL3 CTL fusion partner. Neither partner and none of the hybrid clones lysed the CA14/ca target cells in a standard 3.5 h cytolytic assay, although these target cells are lysed by the cytolytic 6TGOL3 cells in the presence of ConA. Similar results were obtained using lipopolysaccharide (LPS)-stimulated $I-A^k$ -bearing CBA spleen cells cultured with conalbumin or LPS-stimulated, $I-A^b$ -bearing B6 spleen cells as targets (data not shown). The 384.5 anti-T cell receptor hybridoma cells can serve as target cells for the L3 clone⁵ and were effectively lysed by the three hybrid clones (data not shown). However, the F23.1 anti-HTL antigen receptor hybridoma cells are not lysed by the hybrid clones (data not shown). A classical CTL clone, dB45 (ref. 5), which binds the F23.1 monoclonal antibody, does lyse these hybridoma target cells.

Lymphokine release is another important measure of cell activation. Supernatants from unfused partner cells and hybrid cells obtained 24 h after various types of stimulation were tested for the presence of TCGF and IFN- γ (Table 1). The 6TGOL3 fusion partner can be stimulated to release IFN- γ by irradiated, T-depleted DBA/2 spleen cells or by immobilized anti-T cell receptor monoclonal antibody 384.5, as previously shown with the L3 clone^{14,15}. No TCGF activity could be detected in these supernatants. The D10 fusion partner releases TCGF (or BSF-1), but not IFN- γ , in response to stimulation with irradiated, T cell depleted CBA spleen cells plus conalbumin or with immobilized monoclonal antibody F23.1, an antibody reactive with an allotypic determinant on the T cell receptor¹⁶. The three hybrid clones each produced detectable levels of both lymphokines in response to all of the stimuli. Thus lymphokine production is not tightly linked to the specific antigen receptor. Results obtained with hybrids derived from the fusion of 6TGOL3 with HTL clones L2 and HK16 also show that all lymphokines secreted by both fusion partners are released in response to recognition of the antigen receptor contributed by either partner

Table 1 Lymphokine release from hybrid clones after stimulation by irradiated, T-depleted spleen cells or immobilized anti-T-cell receptor monoclonal antibodies

Clone†	Release of IFN- γ * Stimulus‡				Release of TCGF (units per ml)§ Stimulus‡			
	CBA/ca	DBA	IF23.1	I384.5	CBA/ca	DBA	IF23.1	I384.5
6TGOL3	0	1450	0	1390	0	0	0	0
D10	0	0	0	0	48	0	42	0
Hybrid 1	1100	1550	1200	1450	22	16	25	12
Hybrid 2	1020	1200	1350	1100	38	33	33	31
Hybrid 3	1100	1450	1140	1300	36	22	32	20

* IFN- γ is measured in an assay for activation of tumoricidal activity by bone marrow-derived macrophages¹⁹. Briefly, 3×10^5 cells per ml from C57Bl/6 bone marrow were cultured in the presence of colony stimulating factor and 10% horse serum at 37 °C. Eight days later, macrophages were harvested and cultured at a concentration of 5×10^5 per ml with sample supernatant and LPS. After 24 h, macrophages were washed and 4×10^5 per ml ^{51}Cr -labelled P815 target cells were added. Supernatants were collected 20 h later and % specific ^{51}Cr release measured. Activity is expressed as the reciprocal of the dilution of sample which results in 50% specific lysis.

† Replicate wells containing 5×10^4 cloned T cells were cultured with the appropriate stimulus in 96-well plates for 24 h at 37 °C. Supernatants from three replicate wells were collected and pooled for analysis.

‡ Spleen cells were prepared as described in Fig. 2. Anti-T cell receptor monoclonal antibody 384.5 was immobilized to the wells before the addition of cloned T cells by incubating antibody-containing culture supernatant for 4 h at 4 °C then washing three times. Monoclonal antibody F23.1 was added to wells previously coated with sheep anti-mouse immunoglobulin, incubated for 2 h at 4 °C then washed vigorously.

§ TCGF activity was measured using proliferation of the IL-2 dependent line CTLL-2, clone 20^{20,21}. Briefly, 4×10^3 CTLL-20 cells were added to serial threefold dilutions of sample supernatants. After incubation for 20 h, ^3H -thymidine was added for an additional 4 h. Thymidine incorporation was determined and the dilution yielding 50% of maximum d.p.m. was determined. TCGF activity was expressed in units ml^{-1} based upon comparison with the NCI Biological Response Modifiers Program IL-2 reference reagent.

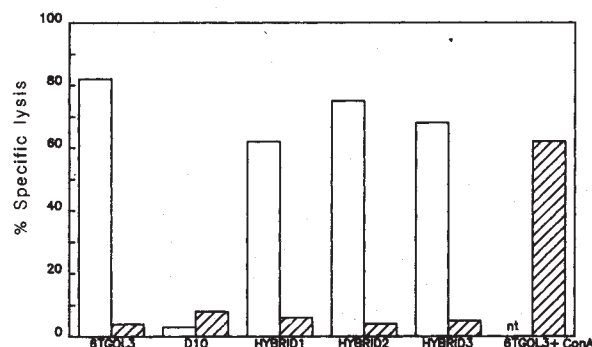


Fig. 3 Cytolysis of: open bars, P815; hatched bars, CA14.11.14 plus conalbumin target cells by hybrid clones. Cytolytic activity was measured as ^{51}Cr release from labelled target cells. CA14.11.14 (I-A^k transfected L cells) were cultured for 24 h with $100 \mu\text{g ml}^{-1}$ conalbumin. These cells and P815 (H-2^d) mastocytes were labelled with ^{51}Cr . 6TGOL3, D10 and three hybrid clones were mixed with the labelled target cells in microtitre trays at an effector to target cell ratio of 10:1 (2.5×10^4 cloned T cells and 2.5×10^3 ^{51}Cr -labelled target cells), centrifuged, then incubated for 3.5 h at 37°C . Supernatants were collected and % specific lysis measured. In a lectin-facilitated assay $20 \mu\text{g ml}^{-1}$ ConA was added to the cytolytic clone L3 and ^{51}Cr CA14.11.14 plus conalbumin targets to verify that these target cells were susceptible to CTL lysis.

in the hybrid T cell clone. Apparently, stimulation of either antigen receptor complex initiates a series of biochemical events leading to transcription of lymphokine genes, expression of which has been programmed by the previous differentiation events in each fusion partner. The failure of CTL clone L3 cells to secrete TCGF does not appear to be the result of a mechanism which represses transcription of the BSF-1 or IL-2 gene. If such a repressor mechanism does operate, the 6TGOL3/D10 hybrid cells would not be expected to secrete BSF-1, assayed in these studies on the basis of its TCGF activity.

Unlike lymphokine production, cytolysis appears to be tightly linked to recognition of antigen by the CTL partner antigen receptor. This linkage was also seen in preliminary studies of hybrids between 6TGOL3 and L2 and HK16. This may be due to a physical linkage between the CTL antigen receptor complex and the lytic machinery; only the binding of antigen to the CTL antigen receptor physically transmits the signal for lysis to the lytic machinery. Staerz *et al.* demonstrated that coupling a monoclonal antibody directed against the T cell antigen receptor to a target cell can make that cell sensitive to lysis by a CTL¹⁷. Our results show that this interaction must occur through the CTL antigen receptor.

Another possibility is that the antigen density on the target cells recognized by the HTL partner antigen receptor is insufficient to induce lysis. T-cell antigen receptor sensitivity to, or affinity for, antigen and the level of specific antigen expression by the target cell are important factors in CTL-mediated cytolysis¹⁸. However, the antigen density was sufficient to trigger lymphokine release and proliferation by the hybrids even though lysis of the HTL antigen-bearing target did not occur. Also, as the F23.1 hybridoma target cells were lysed by the CTL clone dB45, the antigen was sufficiently dense to allow CTL lysis.

We are investigating the effects of anti-Lyt-2 and anti-L3T4 monoclonal antibodies on cytolysis, proliferation, and lymphokine release. These cytolytic-helper hybrid T-cell clones obtained from normal, non-transformed T-cell clones should be useful for obtaining additional information on the link between receptor specificity and functional activities and should provide further insight into the mechanisms and regulatory processes that control T-cell responses.

This work was supported by NIH grant CA-19266 and predoctoral traineeship grant AI-07099-08 from USPHS. We thank C. Janeway for the D10.G4.1 clone and D. Lancki and M. Bevan for the 384.5 and F23.1 hybridoma cell lines. We also thank D.

Dialynas and D. Lancki for helpful discussions and B. Clinchy for technical assistance with lymphokine assays.

Received 11 July; accepted 20 October 1986.

1. Taniguchi, M. & Miller, J. F. A. P. *J. exp. Med.* **148**, 373-382 (1978).
2. Marrack, P. *et al.* in *Isolation, Characterization & Utilization of T Lymphocyte Clones* (eds Fathman, C. G. & Fitch, F. W.) 119-126 (Academic, New York, 1982).
3. Harris, D., MacDonald, H. R. & Cerottini, J.-C. *J. exp. Med.* **159**, 261-275 (1984).
4. Harris, D., MacDonald, H. R. & Cerottini, J.-C. *J. exp. Med.* **159**, 330-335 (1984).
5. Lancki, D. W., Ma, D., Havran, W. L. & Fitch, F. W. *Immun. Rev.* **81**, 65-94 (1984).
6. Glasebrook, A. L. *et al.* *Immun. Rev.* **54**, 225-266 (1981).
7. Kaye, J., Porcelli, S., Tite, J., Jones, B. & Janeway, C. A. *J. exp. Med.* **158**, 836-856 (1983).
8. Prystowsky, M. B. *et al.* *J. Immun.* **129**, 2337-2344 (1982).
9. Herold, K. L., Lancki, D. W., Moldwin, R. L. & Fitch, F. W. *J. Immun.* **136**, 1315-1321 (1986).
10. Tite, J. P., Powell, M. B. & Ruddle, N. H. *J. Immun.* **135**, 25-33 (1985).
11. Kohler, G. & Milstein, C. *Nature* **256**, 495-497 (1975).
12. Harwell, L., Skidmore, B., Marrack, P. & Kappler, J. *J. exp. Med.* **152**, 893-904 (1980).
13. Malissen, B. *et al.* *Cell* **36**, 319-327 (1984).
14. Lancki, D. W., Lorber, M. I., Loken, M. L. & Fitch, F. W. *J. exp. Med.* **157**, 921-935 (1983).
15. Moldwin, R. L., Lancki, D. W., Herold, K. L. & Fitch, F. W. *J. exp. Med.* **163**, 1566-1582 (1986).
16. Staerz, U. D., Rammensee, H.-G., Benedetto, J. D. & Bevan, M. J. *J. Immun.* **134**, 3994-4000 (1985).
17. Staerz, U. D., Kanagawa, O. & Bevan, M. J. *Nature* **314**, 628-631 (1985).
18. Shimonkevitz, R., Luescher, B., Cerottini, J.-C. & MacDonald, H. R. *J. Immun.* **135**, 892-899 (1985).
19. Kelso, A., Glasebrook, A. L., Kanagawa, O. & Brunnen, K. T. *J. Immun.* **129**, 550-556 (1982).
20. Gillis, S., Ferm, M. M., Ou, W. & Smith, K. A. *J. Immun.* **120**, 2027-2031 (1978).
21. Ely, J. *et al.* *J. Immun.* **127**, 2345-2349 (1981).

Herpes simplex virus type-1 can reactivate transcription of latent human immunodeficiency virus

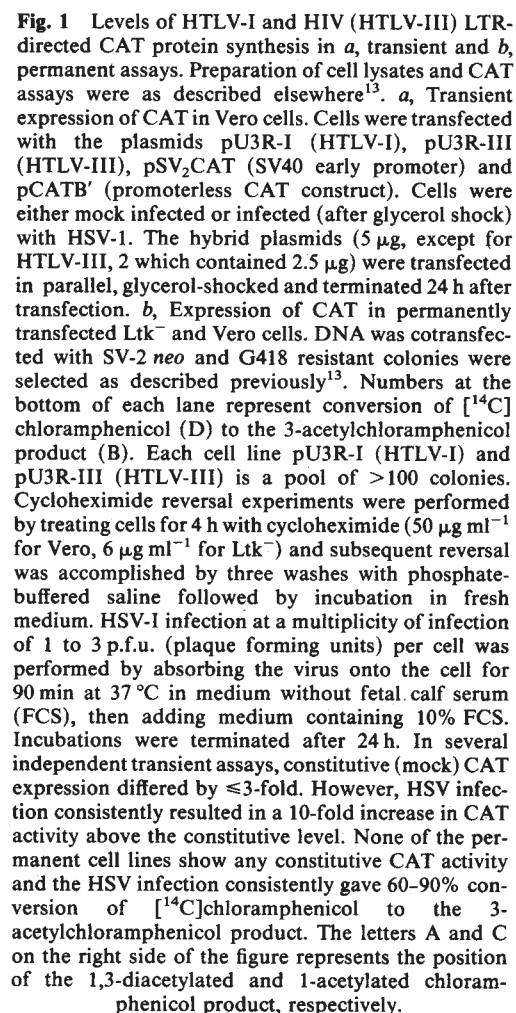
J. D. Mosca*, D. P. Bednarik*, N. B. K. Raj*, C. A. Rosen†, J. G. Sodroski†, W. A. Haseltine† & P. M. Pitha*

* The Johns Hopkins University School of Medicine, Oncology Center, 600 North Wolfe Street, Baltimore, Maryland 21205, USA

† Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA

The human immunodeficiency virus, HIV (formerly T-cell lymphotropic virus type III, HTLV-III or lymphadenopathy-associated virus, LAV) is the primary cause of AIDS (acquired immune deficiency syndrome)². Patients with AIDS have profound immunosuppression as a result of almost complete absence of the OKT4⁺ cell population^{1,3} and are predisposed to a number of opportunistic infections as well as to certain malignant diseases such as Kaposi's sarcoma and B-cell tumours^{4,5}. The majority of the opportunistic infections observed in AIDS patients are from the herpesvirus group⁶⁻⁸ and these are frequently the cause of death in AIDS patients. We have therefore investigated the effect of herpes virus infection on the expression of HIV and we provide evidence that herpes simplex virus type I (HSV-I) can reactivate transcription of latent HIV. In OKT4⁺ human T-cells HIV replicates to high virus titres, resulting in high level expression of viral RNA. This high level of expression has been attributed to virus-associated *trans*-acting factors that increase gene expression, directed by the HIV long terminal repeats (LTR)⁹, post-transcriptionally. In our studies we have tested whether transcription directed by the LTR of HIV is stimulated by HSV-I.

Previous studies investigating the activity of HIV LTR sequences in lymphoid and non-lymphoid cells were carried out in transient expression assays, where the expression of LTR was measured 24-48 h after transfection. In these experiments, the plasmid pU3R-III (ref. 10) was used; it contains nucleotides -453 to +80 of a 3' HIV LTR inserted 5' to the bacterial chloramphenicol acetyltransferase (CAT) gene. We used the pU3R-III plasmid and constructed both murine (Ltk⁻) and simian (Vero) stable DNA transfected cell lines. For comparison, we constructed permanent cell lines (murine and simian) containing the plasmid pU3R-I (ref. 11), which contains



a

Probe Mock HSV 0.5 1 2 4 8 16h

622-
527-
404-
309-
242-

b

Mock CHX 1 2 4h

T7 CAT HTLV-III

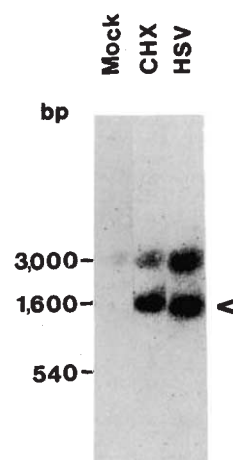
E H A

Probe mRNA 495 nt

S₁

335 bp

Fig. 3 *In vitro* labelling of HIV/CAT RNA transcripts in nuclei isolated from Ltk⁻ cell. Identical results were obtained for Vero cells (data not shown). Isolation of nuclei, conditions for labelling, isolation, and purification of RNA were as described²⁶. The pU3R-III-CAT plasmid DNA was cleaved with *Xho*I, *Bgl*II and *Bam*HI, size separated, denatured, transferred to nitrocellulose, baked, prehybridized and hybridized to RNA labelled with [³²P]UTP *in vitro*. Hybridization was done in a buffer containing 5×SSC, 50% formamide, and 10% dextran sulphate for 3 days at 37 °C; nitrocellulose filters were washed and treated with RNase²⁶. The indicated base pairs (bp) represent the size of the pU3R-III fragments after restriction with *Xho*I, *Bgl*II and *Bam*HI. The 3,000 bp fragment (*Bam*HI to *Xho*I) contains the pBR322 portion of the plasmid, whereas the 1,600 bp (*Bgl*II to *Bam*HI) fragment contains sequences from the HIV LTR cap-site (*Bgl*II), the entire CAT coding region and SV40 polyadenylation site (*Bam*HI) (arrow). The 540 bp (*Xho*I to *Bgl*II) fragment contains the HIV promoter, minus sequences 3' to the mRNA cap-site. Mock, cycloheximide (CHX) and HSV-I treatments as described (Fig. 1).



the U3, R and 105 bp of the U5 regions of a 3' HTLV-I LTR¹², 5' to the CAT gene. In transient assays in Vero cells, pU3R-I (HTLV-I) and pU3R-III (HTLV-III) expressed CAT activity with similar efficiency, but less than that observed for the early promoter of simian virus 40 (SV40) (SV2CAT) (Fig. 1a, Mock). Herpes simplex virus type I (HSV-I) infection increased HIV LTR directed CAT activity in transient assays (Fig. 1a). When the amount of transfected plasmid was decreased, from 5 µg (HTLV-III, 1; Fig. 1a) to 2.5 µg (HTLV-III, 2), the constitutive (Mock) HIV LTR directed CAT activity was reduced, but the extent of increase in CAT activity by HSV infection remained the same (~10-fold). HSV-I infection did not affect either HTLV-I LTR, SV2CAT or pCATB' expression in transient assays. However, once integrated, the HIV LTR, but not the HTLV-I LTR, failed to express CAT constitutively (Fig. 1b, Mock). Experiments on two independently pooled murine and simian cell lines show that expression from the HIV LTR can be repeatedly restored either by treatment with cycloheximide or by infection with HSV-I (Fig. 1b). Cycloheximide treatment does not affect HIV LTR expression in transient assays. Pooled cell lines were used instead of selected colonies to average out any fluctuations observed with individual cell clones^{13,14}, but when individual clones were tested, eleven out of the twelve cell lines were also inducible by HSV-I infection (data not shown). In addition to HSV-I, HSV-2 and cytomegalovirus (CMV) could also reactivate the latent HIV LTR in our cell lines¹⁵. These experiments show that the HIV LTR was inactive when permanently integrated into the host chromatin and that HSV infection could reactivate the HIV LTR both in a latent state and in transient assays.

Two lines of evidence suggest that the HSV and the HIV trans-activating gene product (tat III) reactivate HIV LTR at different levels. First, S₁ analysis showed that there were no constitutive correctly-initiated HIV/CAT messenger RNA transcripts present in our cell lines (Fig. 2, Mock). Second, the run-on transcription in isolated nuclei showed no detectable HIV LTR/CAT transcripts in these cells (Fig. 3, Mock); only after HSV infection was CAT transcription detected (Figs 2a and 3). These data show that the activation by HSV-I infection occurred at the transcriptional level. In addition to HSV infection, HIV LTR expression could be induced by cycloheximide treatment. The HIV/CAT messenger RNA accumulation was seen within 1 h of cycloheximide treatment (Fig. 2b) and run-on transcription analysis indicated that cycloheximide induced the transcription of the HIV/CAT gene (Fig. 3).

To demonstrate the requirement for HSV-I gene expression in the activation of the HIV LTR, we examined the patterns of HIV/CAT mRNA induction after infection with a viral mutant defective in HSV gene expression¹³ (Fig. 4). Infection with HSV-I (HFEM)tsB7 at the non-permissive temperature prevents release of HSV viral DNA into the cellular nucleus¹⁶. Infection

with either tsB7 or the wild-type HSV virus reactivated the HIV LTR and led to the expression of the CAT gene at permissive temperature (33 °C), but not at the non-permissive temperature (39 °C); the induction with wild-type HSV was not impaired at 39 °C. The tsB7 mutant was used here only to show the requirement for productive infection in the reactivation of the latent HIV LTR. HSV reactivation of latent HIV sequences in these cell lines does not require the 65K virion transcriptional transactivator¹⁷.

It is possible that in some HIV-infected individuals, who do not show symptoms of AIDS¹⁸, opportunistic viral infections could contribute to the appearance of the disease. T cells are not the only cells that can be infected by HIV. Both macrophages and subpopulations of B cells which express a T4 epitope can be infected¹⁹⁻²¹. HIV has also been detected in the brains of AIDS patients, in endothelial and mononuclear cells and at low levels in astrocytes and neurones²². Since HSV replicates in neurones and undifferentiated monocytes and macrophages²³, it is not unlikely that HSV infection could contribute to HIV activation both in the central nervous system and in cells of lymphoid origin.

We thank Dr G. S. Hayward for his advice and helpful suggestions. This work was supported by the NCI (Core) grant 5P30CA06973-23 (J.D.M.) and American Cancer Society grant MV208-A (P.M.P.).

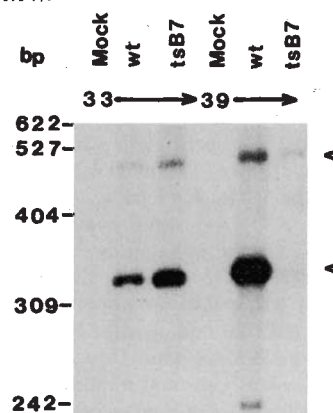


Fig. 4 Requirements for HSV-I gene expression for reactivation of the latent HIV LTR. S₁ analysis of RNA isolated from the permanent Vero cell line containing the pU3R-III plasmid 8 h after infection with wild-type (wt) HSV-1 (MP) or HSV-I(HFEM-STH2)tsB7 at permissive, 33 °C (33) and non-permissive, 39 °C (39). Hybridization and S₁ nuclease digestion were as described (Fig. 2). HSV infection was done in flasks submerged in constant temperature water baths. Arrows show position of the full length probe (top) and the correctly initiated transcript (bottom).

Received 21 July; accepted 15 October 1986.

1. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. *Science* **224**, 497-500 (1984).
2. Barre-Sinoussi, F. *et al.* *Science* **220**, 868-871 (1983).
3. Edwards, K. M., Cooper, M. D., Lawton, A. R., Sanders, D. S. & Wright, P. F. *J. Pediatr.* **105**, 70-72 (1984).
4. Rapp, F. *J. natn. Cancer Inst.* **72**, 783-787 (1984).
5. Friedman-Kien, A. E. *et al.* *Ann. Intern. Med.* **96**, 693-700 (1982).
6. Hirsch, M. S. & Felsenstein, D. *Ann. N.Y. Acad. Sci.* **437**, 8-15 (1984).
7. Spector, S. A., Hirata, K. K. & Neuman, T. R. *J. infect. Dis.* **150**, 953-956 (1984).
8. Quinn, G. V. *et al.* *J. Am. Med. Ass.* **252**, 72-77 (1984).
9. Rosen, C. A. *et al.* *Nature* **319**, 555-559 (1986).
10. Sodroski, J. *et al.* *Science* **227**, 171-173 (1985).
11. Sodroski, J. G., Rosen, C. A. & Haseltine, W. A. *Science* **225**, 381-385 (1984).
12. Rosen, C. A., Sodroski, J. G. & Haseltine, W. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6502-6506 (1985).
13. Mosca, J. D., Reyes, G. R., Pitha, P. M. & Hayward, G. S. *J. Virol.* **56**, 867-878 (1985).
14. Mantei, N., Boll, W. & Weissmann, C. *Nature* **281**, 40-46 (1979).
15. Mosca, J. D. *et al.* (in preparation).
16. Batterson, W., Furlong, D. & Roizman, B. *J. Virol.* **45**, 397-407 (1983).
17. Campbell, M. E. M., Palfreyman, J. W. & Preston, C. W. *J. molec. Biol.* **180**, 1-19 (1985).
18. Blattner, W. A., Biggar, R. J., Weiss, S. H., Melbye, M. & Goedert, J. J. *Ann. intern. Med.* **103**, 665-669 (1985).
19. Dalglish, A. G. *et al.* *Nature* **312**, 763-766 (1984).
20. Montagnier, L. *et al.* *Science* **225**, 63-66 (1984).
21. Koenig, S. *et al.* *Science* **233**, 1089-1093 (1986).
22. Wiley, C. A., Schrier, R. D., Nelson, J. A., Lampert, P. W. & Oldstone, M. B. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7089-7093 (1986).
23. Kirchner, H. *Monogr. Virol.* **13**, 1-104 (1982).
24. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
25. Melton, D. A., Krieg, P. A., Rebayliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. *Nucleic Acids Res.* **12**, 7035-7056 (1984).
26. Raj, N. B. K. & Pitha, P. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3923-3927 (1983).

Bovine papillomavirus E2 trans-activating gene product binds to specific sites in papillomavirus DNA

Elliot J. Androphy*, Douglas R. Lowy & John T. Schiller

Laboratory of Cellular Oncology, Building 37; Room 1B-26, National Cancer Institute, Bethesda, Maryland 20892, USA

Enhancers are *cis*-acting elements that activate transcription in higher eukaryotes independently of their position or orientation relative to the promoter that they activate¹. The mechanisms by which enhancers activate transcription are poorly understood, in part because, with the exception of the glucocorticoid receptor², the proteins that directly interact with enhancers have not been purified, nor have the genes encoding them been cloned³⁻⁵. The upstream regulatory region (URR) that immediately precedes the early genes of the bovine papillomavirus type 1 genome (BPV) has enhancer activity when it is activated by a *trans*-acting gene product of the BPV E2 open reading frame (ORF) (Fig. 1)⁶. It is not known whether this enhancement represents a direct or indirect effect of E2 on the URR. We have used an E2 peptide expressed in bacteria and a DNA-protein complex immunoprecipitation assay to study E2-mediated enhancement of transcription by the URR. We show here that this peptide directly binds to four specific sites in the BPV URR, and to one site in the human papillomavirus (HPV)16 URR. All the binding sites contain a related sequence of nucleotides; a 23 base pair (bp) fragment containing this sequence can specifically prevent binding of the E2 protein to the BPV URR. The BPV E2-URR enhancer interaction may therefore represent a useful model system for studying the mechanism of transcriptional enhancement, as both an effector protein and its target enhancer can be purified and genetically manipulated.

E2 mutants of the full-length BPV genome often have a reduced ability to transform cultured mouse cells morphologically^{7,8} but, unlike some *trans*-activating genes^{9,10}, the E2 gene

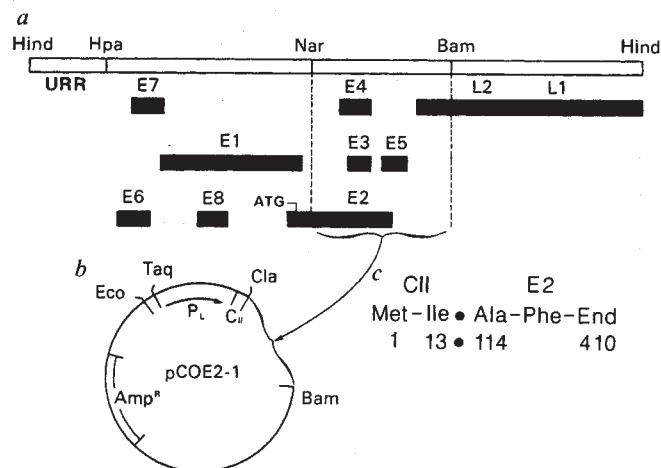


Fig. 1 The 8-kb BPV-1 genome, linearized at the *Hind*III site. The *Hpa*I site is numbered nucleotide 1. *a*, E1-8 are the early region (expressed in cultured cells) open reading frames (ORFs) as deduced from DNA sequence analysis^{15,17}. The L1, L2 ORFs are expressed only in productively infected papillomas. URR is the upstream regulatory region. *b*, To express the E2 ORF, the *Nar*I to *Bam*HI fragment (nucleotides 2,944-4,450) of BPV-1 was inserted into the *Taq*I/*Bam*HI site of pCO-5¹⁸, which provides the P_L promoter, ribosome binding site and 13 amino acids of the phage cII N-terminus. *c*, The resulting vector, pCOE2-1, directs the synthesis of 13 N-terminal amino acids of cII followed in-frame by 297 amino acids of E2 (the first in-frame stop codon is at nucleotide 3838).

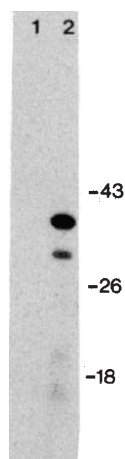
has not been shown to have intrinsic transforming activity. E2 probably affects transformation by positively regulating the expression of the BPV genes E6 and/or E5 that can independently transform cultured cells when activated by heterologous promoters¹¹⁻¹⁴. The conservation of the E2 ORF among the papillomaviruses¹⁵ and the ability of frame-shift mutations in E2⁸ to eliminate its activity imply that the *trans*-acting transcriptional activity of E2 is the result of a translational product of the gene. However, the E2 protein has not yet been identified, presumably because, like other BPV early genes, its transcript is relatively scarce in cultured cells^{16,17}.

To express an E2 peptide in bacteria, the 3' three quarters of the E2 ORF were cloned into pCO-5, an expression vector that contains the phage λ P_L promoter and the amino-terminal end of the cII protein¹⁸ (Fig. 1) to give expression vector pCOE2-1. If translation is initiated at the first ATG of the E2 ORF (nucleotide 2,608) and terminates at the first in-frame stop codon (nucleotide 3,838)¹⁷, expression of pCOE2-1 should result in a 310 amino-acid fusion protein composed of the first 13 amino acids of cII and 297 amino acids of E2 (amino acids 114-410) of relative molecular mass 37,000 (*M_r* 37K). The clone was introduced into an *Escherichia coli* strain that contained the temperature-sensitive λ cI857ts repressor of P_L. After thermal induction, ~2% of total bacterial protein was the predicted 37K E2 fusion protein. The 37K band was extracted from SDS-polyacrylamide gels and used with Freund's adjuvant to immunize rabbits at 3-4-week intervals. The resultant antisera recognized the E2 fusion protein (Fig. 2) but not bacterially synthesized BPV E6 or H-RAS fusion proteins that had the same cII amino termini¹⁸. A band of the same relative molecular mass was also immunoprecipitated with antisera raised against a peptide synthesized *in vitro* derived from the predicted E2 protein sequence, confirming that the 37K band was the E2 fusion protein (data not shown).

We then used the DNA immunoprecipitation assay of McKay¹⁹ (see Fig. 3, legend) to test the ability of the E2 fusion protein to bind specifically to BPV DNA. When end-labelled

* Present address: Department of Dermatology, Tufts-New England Medical Center, Boston, Massachusetts 02111, USA.

Fig. 2 Bacteria, N6405 containing pCOE2-1 were grown in minimal medium, induced at 42 °C for 15 min, and labelled with ^{35}S -methionine for 15 min. After sequential treatment of the bacteria with lysozyme and DNase I, proteins were solubilized with 4 M urea, 1 mM dithiothreitol (DTT). Immunoprecipitations were performed as described¹⁸ with either pre-immune sera (lane 1) or immune sera (lane 2). The 37 K band was seen only with bacteria that contained pCOE2-1. Fusion proteins cII-RAS and cII-E6¹⁸ from the same bacterial expression vector (pCO-5) as used for E2 formed additional controls (not shown because the lanes were as blank as the lane with pre-immune serum).



fragments generated by digestion of the 8 kilobase (kb) BPV genome with a combination of *Bam*HI, *Hind*III and *Sau*96I endonucleases were assayed, two fragments of 498 bp (nucleotides 6,958–7,456) and 233 bp (nucleotides 7,586–7,816) were specifically bound by the protein-antibody complexes (compare lanes 1 and 2 in Fig. 3a). Both these fragments are within the segment of the URR that has enhancer activity. We conclude that the binding to these two fragments results from a sequence-specific interaction with the E2 fusion protein, as no binding was detected when pre-immune sera were used in the assay (lane 4) or when the same cII peptide was linked to the BPV E6 protein or the H-RAS protein and the assay performed with either the E2 antisera or antisera that recognize the E6 or H-RAS peptides¹⁸ (lanes 5–8). Specific binding occurred over a pH

range of 6.8–7.4, a temperature range 4–37 °C and in concentrations of Non-idet P40 of 0.05–0.5% (data not shown).

When the BPV DNA fragments were thermally denatured before incubation with the E2-antibody complexes, all the single strand fragments longer than 100 nucleotides were bound (lane 3), indicating that the E2 protein binds single-stranded DNA non-specifically. Consistent with the observation that the E2 of BPV can *trans*-activate the transcription of the HPV16 early region (L. Turek; W. Phelps and P. Howley, personal communications), the BPV E2 fusion protein also reacted specifically with the HPV16 URR²⁰, by binding to a single 88-bp fragment (nucleotides 24–112) (compare lanes 9 and 10). Neither the simian virus 40 (SV40) genome (compare lanes 11 and 12), nor the Harvey murine sarcoma virus long terminal repeat (not shown), that both contain enhancers^{1,21}, possess sequences that are specifically recognized by the E2 protein in this assay. These results indicate that DNA sequences recognized by E2 are not common to all enhancer elements.

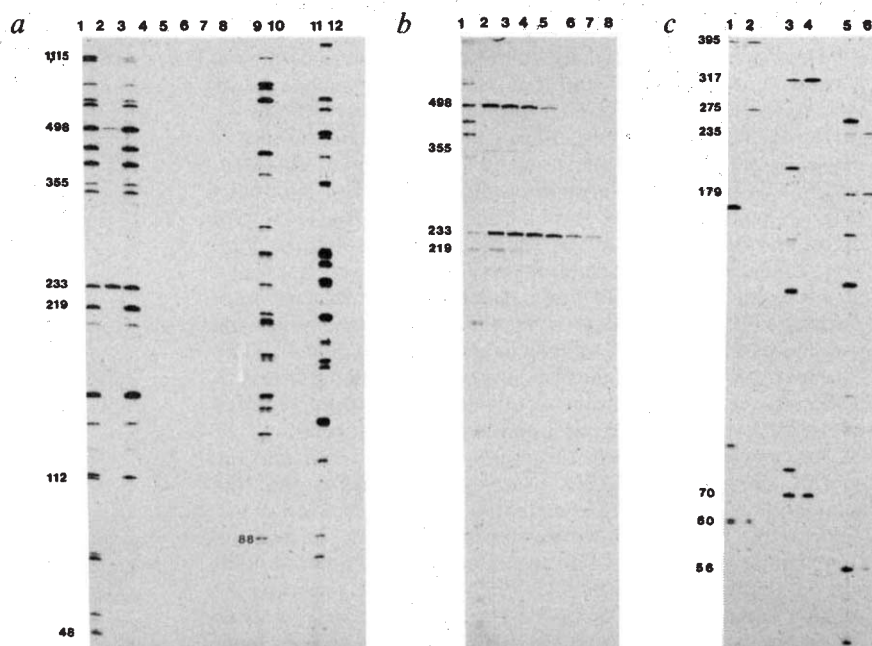
Because other sequence-specific DNA-binding proteins may bind DNA non-specifically when assayed under conditions of excess protein, the above experiments were performed with low concentrations of E2 protein. When increased amounts of E2 were used, two additional binding sites in the BPV genome were revealed (Fig. 3b). One is in a 219-bp fragment that is located further upstream in the URR (nucleotides 7,819–7,893); the other is in a 355-bp fragment located in the E2 ORF (nucleotides 2,904–3,259). We conclude that E2 has a lower affinity for the binding sites in these two fragments than for the binding sites in the fragments described above. At least one additional binding site was also detected in the HPV16 genome when the concentration of E2 protein was increased (data not shown).

To determine more precisely the number and location of the high-affinity E2 protein binding sites, we isolated the 987-bp *Hind*III to *Hpa*I URR fragment (nucleotides 6,958–7,945) and

Fig. 3 Protein-DNA complex immunoprecipitation.

a, Specific binding of E2 protein to DNA; **b**, titration of specific E2 protein binding sites; **c**, localization of high-affinity E2 binding sites. **a**, Lanes 1–8: immunoprecipitation of the 8-kb BPV genome restricted with *Bam*HI, *Hind*III and *Sau*96I. Except for lane 1, which contains end-labelled BPV DNA marker, each lane shows DNA after immunoprecipitation. Lane 2, E2 protein and E2 antisera; lane 3, same as lane 2 except that the DNA was thermally denatured before binding; lane 4, E2 protein and pre-immune sera; lane 5, cII-BPV E6 protein with E2 antisera; lane 6, cII-E6 protein with E6 antisera; lane 7, cII-ras protein with E2 antisera; lane 8, cII-ras protein and anti-ras monoclonal 259. Lanes 9 and 10 show immunoprecipitation of HPV16 DNA digested with *Bam*HI, *Dde*I and *Sau*96I: lane 9, HPV16 DNA marker; lane 10, after immunoprecipitation with E2. Lanes 11 and 12 show immunoprecipitation of SV40/pBR322 DNA digested with *Bam*HI and *Sau*96I. Lane 11, SV40/pBR322 DNA marker; lane 12, after immunoprecipitation with E2. **b**, Lane 1, BPV marker DNA digested with *Bam*HI, *Hind*III and *Sau*96I. In lanes 2–8, the immunoprecipitation assay was performed as described above except that 200, 100, 50, 25, 12.5, 6.25 and 3.1 ng of E2 protein were used, respectively. **c**, Localization of high-affinity E2 binding sites in the BPV URR. Assays were performed with the 987-bp *Hind*III to *Hpa*I fragment instead of the entire BPV genome. Binding was tested after digestion with *Fok*I, lane 1 (marker) and lane 2 (immunoprecipitation); *Dde*I and *Taq*I, lane 3 (marker) and lane 4 (immunoprecipitation); or *Hpa*II and *Sau*96I, lane 5 (marker) and lane 6 (immuno-precipitation). The 235-nucleotide fragment seen in the immunoprecipitation of the *Hpa*II-*Sau*96I fragments is the result of incomplete digestion by *Hpa*II of the contiguous 179-bp and 56-bp fragments that individually bind E2.

Methods. Aliquots (50 ng) of partially purified E2 protein (50% by acrylamide gel analysis) were diluted with DIB (20 mM HEPES, pH 7.2, 150 mM KCl, 0.05% Non-idet P-40, 1 mM EDTA, 1 mM DTT, 1% aprotinin) and incubated at 4 °C with E2-specific antisera. Complexes were collected with protein A-Sepharose and washed with DIB. After restriction endonuclease digestion and end-labelling with [^{32}P]dNTPs by the Klenow fragment of DNA polymerase I, 20 ng of DNA were added in 0.2 ml of DIB containing 400 ng of unlabelled pML2d²¹. After 1 h at 37 °C, the complexes were pelleted and washed 4 times with DIB, dissociated in 1% SDS at 65 °C, phenol-chloroform extracted and the released DNA ethanol-precipitated after the addition of carrier DNA. The resuspended DNA was denatured and analysed on standard acrylamide sequencing gels or, for panel **a**, on a gradient sequencing gel. Gels were dried before autoradiography. When the labelled DNAs were run as relative molecular mass markers, 20% of the amount used in the binding assay was loaded onto the gel. The lengths of representative fragments are indicated (in nucleotides) to the left of each panel.



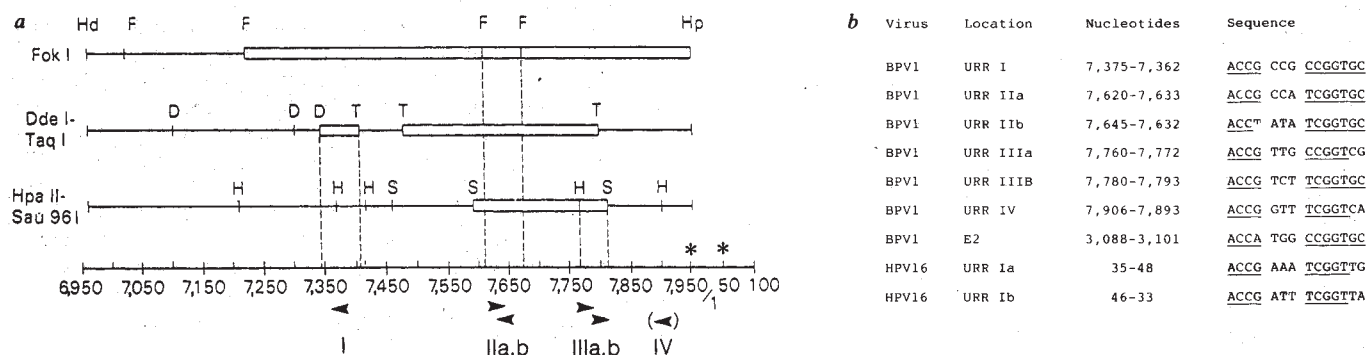


Fig. 4 Localization of E2 binding sites within the BPV URR. *a*, Linear map of the BPV URR; asterisks represent potential TATA transcription signals. The restriction maps generated by the endonucleases listed on the left are drawn above and summarize the data from Fig. 3c. Open boxes, DNA fragments specifically bound by E2; solid lines, fragments not recognized by E2. Arrowheads below the genome map, related motifs (see *b*) that may be associated with E2-protein binding. The arrows point 5' to 3' along the strand of DNA on which the sequence is located, the motifs being present on both strands of the DNA. The parentheses surround the motif in the low affinity binding fragment. Restriction sites: Hd, *Hind*III; Hp, *Hpa*I; F, *Fok*I; D, *Dde*I; T, *Taq*I; H, *Hpa*II; and S, *Sau*96I. *b*, Consensus sequence in the E2 binding sites. The observed E2-protein recognition sites were compared for related sequences. The location of the sites refers to the motifs in *a* and the common oligonucleotides are shown as they would appear 5'-3'.

tested the ability of E2 to recognize this segment after its digestion with various restriction endonucleases (Figs 3 and 4). Three fragments of 395, 275 and 60 bp were immunoprecipitated after *Fok*I digestion (Fig. 3c, compare lanes 1 and 2). Digestion with *Dde*I and *Taq*I revealed two binding fragments of 317 and 70 bp (compare lanes 3 and 4) and digestion with *Hpa*II and *Sau*96I generated two binding fragments of 179 and 56 bp (compare lanes 5 and 6). Because the 56, 60 and 70 bp fragments are non-overlapping, the E2 protein must recognize at least three high-affinity elements in the URR that are located between nucleotides 7,366-7,406, 7,624-7,683 and 7,767-7,822 (Fig. 4).

The DNA sequences of the fragments to which the protein bound were compared to see if they contained related sequences that might be responsible for recognition by the E2 protein. All the fragments that specifically bound the E2 protein contained a similar motif which had the consensus sequence of 5' ACC(G)NNNPYCGGT(GC) 3' (nucleotides in parentheses are preferred but not invariant, N can be any nucleotide and Py can be C or T). This motif is found on either strand of the DNA, and in two instances two copies are found in close proximity (Fig. 4). Note that site I is not bound in the *Hpa*II-*Sau*96I digest and that *Hpa*II cleaves the putative E2-protein recognition sequence. Motifs resembling these elements were not found in the segments of the BPV and HPV16 genome that were not recognized by E2 protein, nor were they found in the other viral genomes, including SV40, polyoma, bovine leukaemia virus and Moloney murine leukaemia virus surveyed in a computer search. Strikingly, sequences similar to this motif have been reported in the URRs of all other papillomaviruses sequenced²².

The correlation between the presence of this motif and the specific binding of the E2 fusion protein implies that the sequence-specific binding is mediated by this conserved motif. To test this hypothesis, we examined the ability of a 23-bp fragment encoding motif IIIb to inhibit the specific binding of the E2 peptide to BPV DNA. When incubated with the E2 peptide and the BPV DNA, this fragment effectively blocked the immunoprecipitation of the BPV fragments (Fig. 5, lanes 2-5); it was ~1,000 times more effective than another 23-bp fragment from the URR that did not contain the motif (lanes 6-9). From these results we conclude that this motif, when flanked by several additional base pairs, is sufficient for specific E2-protein binding. Neither the single-strand oligonucleotide containing the motif nor its complement blocked the immunoprecipitation of the double-stranded BPV fragments (data not shown), indicating that this sequence is specifically recognized by E2 protein only in the form of double-stranded DNA.

The experiments reported here imply that the specific binding of E2 peptide to the URR is important for its transcriptional *trans*-activation activity, although we have not yet demonstrated the association *in vivo* of an E2 protein with BPV DNA. Attempts to produce meaningful DNase I footprints have not been successful, largely because the E2 fusion protein is poorly soluble in the absence of urea, and DNase I protection experiments require urea-free conditions. The immunoprecipitation assays were routinely performed under stringent conditions (0.15 M KCl, pH 7.2 at 37 °C), indicating a strong association of the protein with DNA. Under these conditions the E2 protein binds almost exclusively to highly specific sites of the BPV genome. Four of the five binding sites are within the BPV URR and at least one site in the URR of HPV16 is also bound, a result consistent with the previous observation that the URR is the target of E2 activity⁶. The fact that the motif is found more than

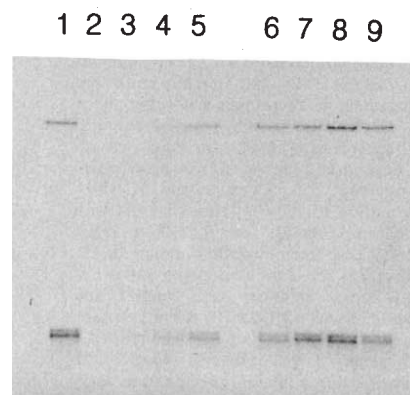


Fig. 5 Competitive inhibition of E2 protein-BPV DNA immunoprecipitation by a 23-bp fragment encoding the presumptive E2 binding site motif. Assays were performed with 25 ng of E2 protein as described in Fig. 3 except that 1-1,000 ng of a competing unlabelled fragment were added to the labelled BPV DNA before its incubation with the antibody-E2 complexes. Lane 1: no additional DNA added; lanes 2-5: 1,000, 100, 10 and 1 ng, respectively, of the double-stranded fragment representing nucleotides 7,774-7,796 (GGTCAAACCGTCTTCGGTGCTC; the conserved motif is underlined) were added; lanes 6-9: 1,000, 100, 10 and 1 ng, respectively, of the double-stranded fragment representing nucleotides 7,698-7,721 (GCGCATAATCAGCTTAATTGGTG, which does not contain the conserved motif) were added. The upper band is the 498-bp binding fragment and the lower band is the 233-bp binding fragment.

once in the BPV URR, and in the URRs of HPVs as well, but was not detected in the genomes of the other viruses screened, supports the suggestion that they may have biological functions. A bacterial fusion protein with a short peptide (74 amino acids) of the *trans*-activating GAL4 protein of yeast binds the same DNA sequences as the full-length 881 amino-acid protein²³, indicating that biologically relevant DNA binding can be mediated by a protein domain that is significantly smaller than the E2 polypeptide present in the E2 fusion protein.

Our results suggest that the E2-protein binding sites detected *in vitro* may represent important components of the enhancer; this would indicate that it may be structurally similar to other enhancers that have been shown genetically to contain multiple and repetitive functional elements^{24,25}. This interpretation is supported by recent genetic data that localize E2-protein-dependent enhancement to segments of the URR that include these motifs (L. Turek; B. Spalholz, P. Lambert and P. Howley, personal communications). Mutant URRs that lack all the motifs (I-IV) do not enhance transcription from the BPV early promoter. URR segments that retain only IIIa, IIIb and IV possess significant enhancing activity, whereas those that also contain IIa and IIb have even greater activity.

Our results localize the sequence-specific DNA-binding activity of the putative E2 protein to the carboxy-terminal three quarters. Whether the amino-terminal amino acids of the ORF are also required for *trans*-activation activity remains to be determined. Perhaps, as with the GAL4 protein²³, a single domain can confer specific DNA binding with other regions of the protein required for *trans*-activation.

In contrast to E2 protein, the best characterized DNA-binding proteins that participate in transcription in higher eukaryotes, such as Sp1²⁶ and USF²⁷, do not appear to interact with enhancer elements. Furthermore, other viral *trans*-acting transcriptional activators, such as adenovirus E1a gene product²⁸ and the HTLV I and II *tat* gene products^{29,30}, may activate transcription indirectly, considering they have not been shown to have DNA-binding activity. If, as we expect, the E2 protein stimulates transcription by directly binding to the URR enhancer, detailed genetic and biochemical studies of the mechanism of enhancement could be undertaken in cultured cells and cell-free extracts using the BPV system, because an effector protein, its gene, its target enhancer and specific antisera are all available.

In addition to their relevance to the study of enhancement, our observations imply that E2 protein may stimulate BPV transformation *in vitro* and participate in the viral life cycle by binding to specific sequences of the viral DNA (and possibly to analogous sequences in cell DNA). It might be possible to develop pharmacological agents that interfere with viral infection by inhibiting E2 protein-DNA interactions.

22. Dartmann, K., Schwarz, E., Gissmann, L. & zur Hausen, H. *Virology* **151**, 124-130 (1986).
23. Keegan, L., Gill, G. & Ptashne, M. *Science* **231**, 699-704 (1986).
24. Herbolme, P., Bourachot, B. & Yaniv, M. *Cell* **39**, 653-662 (1984).
25. Zenke, M. *et al.* *EMBO J.* **5**, 387-397 (1986).
26. Dynan, W. S. & Tjian, R. *Nature* **316**, 774-778 (1985).
27. Sawadogo, M. & Roeder, R. *Cell* **43**, 165-175 (1985).
28. Kovsed, I., Reichel, R. & Nevins, J. *Cell* **45**, 219-228 (1986).
29. Felber, B. K., Paskalis, H., Klienman-Ewing, C., Wong-Staal, F. & Pavlakis, G. N. *Science* **229**, 675-679 (1985).
30. Rosen, C. A., Sodroski, J. G., Kettman, R. & Haseltine, W. A. *J. Virol.* **57**, 738-744 (1986).
31. Lusky, M. L. & Botchan, M. R. *Nature* **293**, 79-81 (1981).

Evidence against Ha-ras-1 involvement in sporadic and familial melanoma

Daniela S. Gerhard*†, Nicholas C. Dracopoli*‡, Sherri J. Bale§, Alan N. Houghton||, Paul Watkins¶, Christine E. Payne*, Mark H. Greene§ & David E. Housman*

* Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

‡ Cancer Immunology Laboratory, Sloan-Kettering Institute, New York, New York 10021, USA

§ Environmental Epidemiology Branch, National Cancer Institute, Bethesda, Maryland 20205, USA

|| Solid Tumor Immunology Laboratory, Sloan-Kettering Institute, New York, New York 10021, USA

¶ Integrated Genetics, Framingham, Massachusetts 01710, USA

It was recently reported that different rare alleles at the Ha-ras-1 locus occurred at a significantly higher combined frequency in cancer patients than in an unaffected population¹. In particular, melanoma patients were reported to have a significantly higher frequency of such alleles. We have examined the frequency of rare Ha-ras-1 alleles in a large number of cases of sporadic melanoma. Our results indicate that the distribution of rare alleles in this population does not differ from that found in normal populations. Also, to test the hypothesis that a hereditary predisposition to melanoma could be inherited via an allele at the Ha-ras-1 locus, we examined the transmission of the segment of the short arm of chromosome 11 (11p) carrying the Ha-ras-1 locus in a number of families previously shown to exhibit a hereditary predisposition to melanoma and its precursor lesion, the dysplastic nevus syndrome (DNS)^{2,3}. Our genetic linkage results thus obtained strongly exclude the association of a predisposition to melanoma or the precursor lesion with the inheritance of the Ha-ras-1 locus or the segment of chromosome 11 on which it is located. These results imply that hereditary predisposition to melanoma is associated with genes other than the Ha-ras-1 locus, contradicting the original suggestion of Kroutiris *et al.*¹, made on the basis of either an inadequate sample size or other misleading experimental factors.

We examined DNA samples from 58 unrelated patients with sporadic malignant melanoma to determine if a high frequency of variant Ha-ras-1 alleles could be identified in these patients. A control population of 43 unrelated individuals of similar age, sex and ethnic background was also examined. DNA samples extracted from white blood cells or Epstein-Bar virus-transformed B-cell lines were sequentially digested with restriction enzymes *MspI* and *HpaII* and fractionated in 1.4% agarose gels. Eight different Ha-ras-1 alleles were identified (Fig. 1). The four common alleles A1 (1.2 kilobases; kb), A2 (1.6 kb), A3 (2.1 kb) and A4 (2.6 kb) occurred with similar combined frequencies in melanoma patients (0.96) and the normal population sample (0.94) (Table 1). Four rare alleles were observed: the A5 (1.3 kb) and A6 (1.25 kb) alleles were observed in both populations, and two additional rare alleles, A7 (1.15 kb) and A8 (2.2 kb), were observed only in the normal population

Received 9 September; accepted 20 October 1986.

1. Khoury, G. & Gruss, P. *Cell* **33**, 313-314 (1983).
2. Weinberger, C. *et al.* *Science* **228**, 740-742 (1985).
3. Bohnlein, E. & Gruss, P. *Molec. cell Biol.* **6**, 1401-1411 (1986).
4. Singh, H., Sen, R., Baltimore, D. & Sharp, P. *Nature* **319**, 154-158 (1986).
5. Wildeman, A. G. *et al.* *Molec. cell Biol.* **6**, 2098-2105 (1986).
6. Spalholz, B., Yang, Y.-C. & Howley, P. M. *Cell* **42**, 183-191 (1985).
7. Sarver, N., Rabson, M. S., Yang, Y.-C., Byrne, J. C. & Howley, P. M. *J. Virol.* **52**, 377-388 (1984).
8. DiMaio, D. *J. Virol.* **57**, 475-480 (1986).
9. Keller, J. M. & Alwine, J. C. *Cell* **36**, 381-389 (1984).
10. Setoyama, C., Frunzio, R., Liau, G., Mudryj, M. & de Crombrughe, B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3213-3217 (1986).
11. Schiller, J. T., Vass, W. C. & Lowy, D. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7880-7884 (1984).
12. Yang, Y.-C., Spalholz, B. A., Rabson, M. S. & Howley, P. M. *Nature* **318**, 575-577 (1985).
13. Schiller, J. T., Vass, W. C., Voudsen, K. H. & Lowy, D. R. *J. Virol.* **57**, 1-6 (1986).
14. Groff, D. E. & Lancaster, W. D. *Virology* **150**, 221-230 (1986).
15. Daños, O., Giri, I., Thierry, F. & Yaniv, M. *J. invest. Derm.* **83**, 8s-11s (1984).
16. Heilman, C., Engel, L., Lowy, D. R. & Howley, P. M. *Virology* **119**, 22-34 (1982).
17. Stenlund, A., Zabielski, J., Ahola, H., Moreno-Lopez, J. & Pettersson, U. *J. molec. Biol.* **182**, 541-554 (1985).
18. Androphy, E. J., Schiller, J. T. & Lowy, D. R. *Science* **230**, 442-445 (1985).
19. McKay, R. D. G. *J. molec. Biol.* **145**, 471-488 (1981).
20. Seedorf, K., Krammer, G., Durst, M., Suhai, S. & Roweckamp, W. G. *Virology* **145**, 181-185 (1985).
21. Kriegler, M. & Botchan, M. *Molec. cell Biol.* **3**, 325-339 (1983).

† Present address: Department of Genetics, Washington University, St Louis, Missouri 63110, USA.

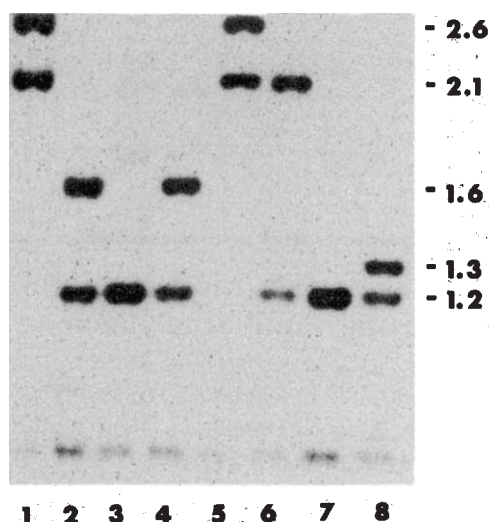


Fig. 1 Southern blot analysis of peripheral white blood cell DNA hybridized with nick-translated plasmid DNA containing the human *Ha-ras-1* gene^{10,11}. The fragment sizes (in kilobases) are on the right. The genotypes are as follows: lane 1, A3/A4; lane 2, A1/A2; lane 3, A1/A1; lane 4, A1/A2; lane 5, A3/A4; lane 6, A1/A3; lane 7, A1/A1; lane 8, A1/A5.

Method. The extraction of genomic DNA from peripheral white blood cells isolated from 10 ml whole blood samples was done by previously described techniques¹². DNA (10 µg) was sequentially digested with *MspI* and *HpaII* and separated in 1.4% horizontal agarose gels and blotted onto Nytran filters (Schleicher and Schuell). The filters were baked for 2 h at 80 °C and prehybridized at 42 °C overnight in 6×SSC, 1×Denhardt's, 1% SDS and 50 µg ml⁻¹ denatured salmon sperm DNA. The filters were then hybridized for 24 h at 42 °C in 50% formamide, 6×SSC, 1% SDS, 50 µg ml⁻¹ denatured salmon sperm DNA and 1×10⁶ c.p.m./per ml of labelled pT22 plasmid DNA (ref. 9), which had been nick-translated to specific activity of >1×10⁸ c.p.m. µg⁻¹. The filters were sequentially washed in 6×SSC, 0.1% SDS (30 min at room temperature); 1×SSC, 1% SDS (30 min at 37 °C); and 0.1×SSC, 1% SDS (1 h at 65 °C), and exposed to Kodak XAR-5 X-ray film at -70 °C for 2-5 days using two Dupont Lightning-Plus intensifying screens.

sample. Rare alleles were found only in heterozygotes with one of the common alleles in both populations. The combined frequency of rare alleles was 0.04 in the sample of melanoma patients and 0.06 in the control population. Krontiris *et al.*¹ observed a combined frequency of rare *Ha-ras-1* alleles of 0.11 in cancer patients. The frequency of rare alleles at the *Ha-ras-1* locus in the sporadic malignant melanoma sample described here is not significantly different from the control population

Table 1 Comparison of *Ha-ras-1* allele frequencies of patients with sporadic melanoma with a normal population

Ha-ras-1 allele	Size (kb)	Allele frequency	
		Melanoma (n=58)	Normal (n=43)
A1	1.2	0.56	0.57
A2	1.6	0.20	0.12
A3	2.1	0.08	0.10
A4	2.6	0.12	0.15
Total common alleles:		0.96	0.94
A5	1.3	0.03	0.01
A6	1.25	0.01	0.01
A7	1.15	0.00	0.03
A8	2.2	0.00	0.01
Total rare alleles:		0.04	0.06

($P < 0.05$) and is less than half that found by Krontiris *et al.*¹ in cancer patients. Consequently, the presence of variant alleles at the *Ha-ras-1* locus does not appear to be an informative marker for the development of malignant melanoma.

We also performed family studies to assess further the possible role of *Ha-ras-1* alleles in the aetiology of melanoma. DNS, the precursor lesion and melanoma are determined by closely linked genes or possibly by the same gene³. The melanoma/DNS gene is inherited in an autosomal dominant manner with high penetrance. Nearly all members of families who develop melanoma previously exhibit the precursor lesion. We performed segregation analysis of restriction fragment length polymorphisms (RFLPs) of four loci on the chromosomal segment 11p to evaluate the likelihood that the melanoma/DNS gene lies at or close to the *Ha-ras-1* locus.

Seven families segregating for melanoma/DNS were used for a two-point linkage analysis between melanoma/DNS and *Ha-ras-1*. The families contained 95 individuals, of whom 67 were typed for *Ha-ras-1*, thirty-two were affected with melanoma or both melanoma and DNS, and nine had only DNS. Linkage analysis was performed using the computer program LIPED⁴ with the assumptions of life-time penetrance of 99% and a mean age of onset of 30 years (± 12 years)⁵. An estimate of the frequency of the melanoma gene can be derived from the incidence of melanoma in the United States population³. A gene present at a frequency of 0.003 would be observed if all cases of melanoma were due to an autosomal dominant genetic mechanism. Because it is difficult to determine the precise ratio of sporadic to familial melanoma cases, we have performed two-point linkage analysis at several values of the gene frequency for the proposed melanoma gene. Gene frequencies of 0.01 and 0.001 are likely to be overestimates of the actual frequency of the melanoma gene in the population, whereas a population frequency of 0.0001 is likely to be an underestimate. The logs of the odds ratio of linkage to non-linkage at given values of θ (lod scores) obtained are highly consistent at all three gene frequencies (Table 2). Negative lod scores were obtained for

Table 2 Two-point lod scores between melanoma/DNS and *Ha-ras-1*

Assumed gene frequency	θ					
	0	0.1	0.2	0.3	0.4	0.5
0.01	-15.89	-5.29	-2.48	-1.05	-0.30	0.0
0.001	-12.87	-5.02	-2.52	-1.19	-0.42	0.0
0.0001	-24.88	-6.06	-2.71	-1.09	-0.28	0.0

Data from seven families. θ is the recombination fraction.

each of the seven families. A lod score of -12.87 at $\theta = 0$ was obtained for linkage between melanoma/DNS and *Ha-ras-1*. Approximately 24 cM on either side of the *Ha-ras-1* locus can be excluded, with a lod score below -2.00, as the potential site of the melanoma/DNS gene (Table 2).

Multi-point analysis allows more efficient use of family data by enabling the segregation of several loci to be observed simultaneously. A multi-point linkage study was performed using the LINKMAP routine of the computer program LINKAGE⁵, with the information on 51 individuals typed for *Ha-ras-1*, of whom 38 were also typed for RFLPs at loci for calcitonin, the β -globin gene complex and insulin. Disease allele frequency was the same as for the two-point analysis. Using a cumulative normal age of onset distribution determined by the above parameters, three liability classes were defined (penetrance to age 24 = 25%; to age 53 = 50%; 54 and older = 99%). Haldane's mapping function⁶ was used to convert recombination frequencies between loci into map distances. The distances between loci on chromosome 11p were held constant as follows: calcitonin-12 cM- β -globin gene complex-16 cM-insulin-3 cM-*Ha-ras-1* (ref. 7). The position of the melanoma/DNS locus was then varied over the chromosome so that free recombination with the 11p linkage

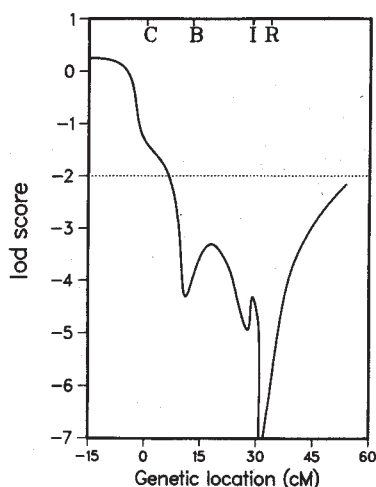


Fig. 2 Likelihood map showing the odds ratio ($10^{\text{lod score}}$) for the location of melanoma/DNS with respect to four linked loci (C, calcitonin; B, β -globin gene complex; I, insulin; R, Ha-ras-1) on 11p, whose distances from each other are fixed. Genetic location in centimorgans (cM) is calculated as distance from calcitonin, the arbitrary origin. The melanoma/DNS locus is shown to be excluded from 50 cM on 11p with odds of 100:1 or greater against its location in this region. The odds ratio reaches $10^{-7.3}$ at a distance of 31 cM from the origin, where Ha-ras-1 is fixed. The map is calculated assuming no sex difference in male and female recombination rates.

group was considered as well as positions between each pair of markers. The relative likelihood of each position compared to the likelihood of melanoma/DNS being unlinked to the 11p markers was computed and converted into a lod score. The results of this analysis showed that the melanoma/DNS locus could be excluded from 50 cM on chromosome 11p, the region 5 cM proximal to β -globin gene complex to 26 cM distal to Ha-ras-1 (Fig. 2).

The two lines of investigation reported here strongly argue against a role for Ha-ras-1 in the aetiology of melanoma. A similar analysis of Ha-ras-1 alleles in patients with myelodysplasia has also failed to confirm an association between malignancy and the presence of rare alleles at this locus⁹. Albino *et al.*¹⁰ have also presented data which argue against a role for the Ha-ras-1 gene in malignant melanoma; they observed only one activated Ha-ras-1 allele among 40 melanoma cell lines tested. Given the exclusion of the Ha-ras-1 region of chromosome 11, other portions of the genome must be tested for linkage with the melanoma/DNS gene to identify the correct location of this gene.

We thank Jen Cunningham for typing the manuscript; B. Nelkin, G. Bell, R. Weinberg, M. Wigler and H. Kazazian for probes; and T. Krontiris for making available Ha-ras-1 results from a family used in this study. This work was supported by NIH grants GM27882 and CA40842 to D.E.H. CA44176 to N.C.D., CA34079 to A.N.H. and Damon Runyon-Walter Winchell Fellowship DRG-665 to DSG.

Received 18 August; accepted 6 October 1986.

1. Krontiris, T. G., Di Martino, N. A., Colb, M. & Parkinson, D. R. *Nature* **313**, 369–374 (1985).
2. Greene, M. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 6071–6075 (1983).
3. Bale, S. J., Chakravarti, A. & Greene, M. H. *Am. J. hum. Genet.* **38**, 188–196 (1986).
4. Ott, J. *Am. J. hum. Genet.* **28**, 528–529 (1976).
5. Lathrop, G. M., Lalouel, J. M., Julier, C. & Ott, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3443–3446 (1984).
6. Haldane, J. B. S. *J. Genet.* **8**, 299–309 (1919).
7. Bale, S. J., Harris, E. L. & Bale, A. E. *Genet. Epidemiol.* (in the press).
8. Albino, A. P., Le Strange, R., Oliff, A. I., Furth, M. E. & Old, L. J. *Nature* **308**, 69–71 (1984).
9. Thein, S. L., Oscier, D. G., Flint, J. & Wainscoat, J. S. *Nature* **321**, 84–85 (1986).
10. Taparowsky, E. *et al. Nature* **300**, 762–765 (1982).
11. Shih, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 5714–5718 (1979).
12. Lowy, D. R., *Proc. natn. Acad. Sci. U.S.A.* **75**, 5539–5543 (1979).

Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor

Stephen Green & Pierre Chambon*

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médecine, 11, rue Humann, 67085 Strasbourg Cédex, France

Steroid hormone receptors are a class of cell-specific *trans*-acting transcription regulatory factors whose activity is controlled by specific binding of the hormone. The hormone-receptor complex appears to associate with promoter/enhancer elements of specific target genes, resulting in activation of transcription (see refs 1 and 2 for reviews). Sequence comparison between the oestrogen^{3,4}, glucocorticoid^{5,6} and progesterone receptors (refs 7, 8 and unpublished results) and site-directed mutation analysis^{9–11}, has identified in each at least two functional domains important for steroid receptor function. Region E (Fig. 1a), is the hormone-binding domain⁹; region C is a 66-amino-acid region (Figs 1a,b) that is more highly conserved than the hormone-binding domain and has the potential to form at least two zinc-stabilized 'DNA-binding fingers' analogous to those proposed for the *Xenopus* transcription factor TFIID^{7,12,13}. We^{3,4} and others¹⁴ have suggested that this region may be the receptor's DNA-binding domain. We show here that point mutations replacing two cysteines by two histidines in the first potential DNA-binding finger of the human oestrogen receptor (hER) prevent it from activating gene transcription. We further show that a chimaeric receptor formed by replacing this 66-amino-acid region of the hER with that of the human glucocorticoid receptor (hGR) activates expression of a glucocorticoid-inducible gene, but not of an oestrogen-inducible gene, in the presence of oestradiol. Thus, region C determines the receptor's specificity for target genes.

Region C of the hER is entirely conserved between chicken and man⁴, is not involved in hormone-binding⁹ and its integrity is required for tight nuclear binding of the hER in the presence of oestradiol⁹. Furthermore, we have shown that region C is essential for the hER to activate transcription¹⁰ of the *vit-tk*-CAT reporter gene, which contains the oestrogen-responsive element (ERE) of the *Xenopus* vitellogenin A2 (ref. 15 and see Fig. 3c). The DNA-binding fingers proposed for region 'C' involve two pairs of cysteines rather than one pair of cysteines and one pair of histidines⁷. We speculated that replacing a pair of cysteines with a pair of histidines might result in a receptor still capable of binding to DNA and of activating transcription. Using site-directed mutagenesis, we chose to replace Cys²⁰² and Cys²⁰⁵ (see Fig. 1b) of the hER with histidines to produce HE27. Transient expression of HE27 in HeLa cells resulted in the production of an oestradiol-binding protein with the same size as the wild-type receptor HEO (Fig. 3a, lanes 2 and 5, and data not shown). No stimulation of CAT activity was observed (not shown), when HE27 (1 μ g) was co-transfected with the reporter gene, *vit-tk*-CAT, into HeLa cells, using conditions identical to those where HEO was stimulatory (see below). This result, which indicates that these two highly conserved cysteines⁷ in this putative DNA-binding finger cannot be functionally substituted with histidines, further demonstrates the importance of region C for activation of transcription.

We reasoned that if region C is responsible for the specific interaction between the receptor and the ERE of the target gene, then the exchange of this region of the hER with that of the hGR should result in a chimaeric receptor able to activate transcription of glucocorticoid-responsive genes. Note that HEO activates transcription of the ERE-containing *vit-tk*-CAT

* To whom correspondence should be addressed.

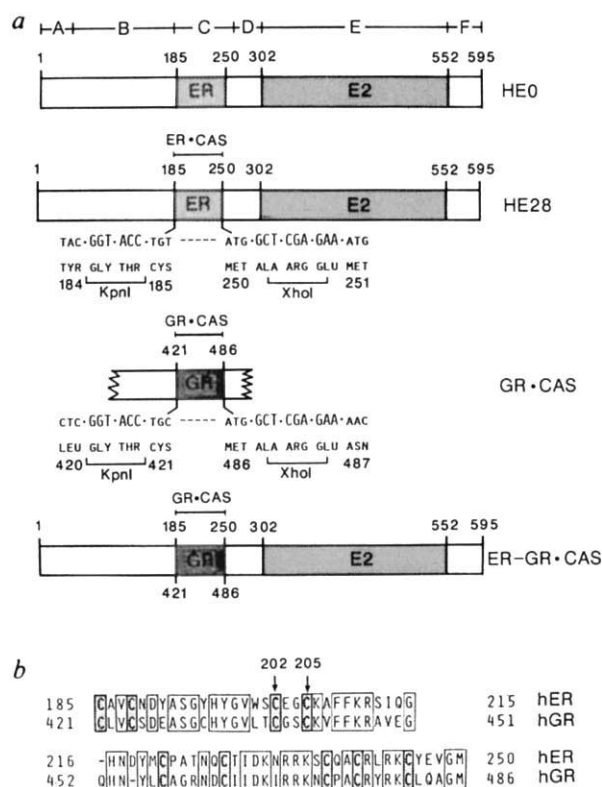


Fig. 1 **a** Construction of the chimaeric receptor ER-GR.CAS. The structures of HEO, HE28, GR.CAS and ER-GR.CAS are shown schematically; the numbers refer to amino-acid positions. The division of wild-type hER (HEO) into regions A-F⁴ is shown at the top of the figure. The putative DNA-binding domain within region C of the hER (ER) and hGR (GR) together with the hER oestrogen-binding domain (E2), are shaded. The *KpnI* and *XhoI* sites, created within region C of the hER and hGR to allow the exchange of this region as a cassette (ER.CAS and GR.CAS), are indicated, with the additional amino acids introduced into the HE28 protein. **b**, Alignment of the hER and hGR within the highly conserved portion of region C⁴. All identical amino acids are boxed and the conserved cysteines shaded. Arrows, hER cysteines (202 and 205) that were replaced with histidines in HE27 (see text). **Methods.** Both the cDNA insert of HEO³ and of an hGR cDNA fragment containing region C (ref. 21 and Govindan, M. Devic & U. Stropp, unpublished results) were inserted into Bluescribe M13+ (Vector Cloning Systems). Mutagenesis of both the hER and hGR sequences was performed as described previously⁹. The oligonucleotides used for the creation of the *KpnI* and *XhoI* sites within the hER and hGR were 30 and 33 nucleotides in length respectively with 12 nucleotides on each side of the new site being complementary to the cDNA sequence. Both the *KpnI* and *XhoI* oligonucleotides were used together during mutagenesis and mutants positive with both oligonucleotide probes were selected. HE27 was created by mutation of HEO using the oligonucleotide primer 5'-GGAGTCTGGTCTCACGAGGGCCACAAGGCC-TTC-3' using techniques already described⁹. This mutation results in a loss of an *AvaII* site (GGTCC to GGTCT) without altering the amino-acid sequence of hER other than the replacement of cysteines 202 and 205 with histidines. Each of the mutant receptor cDNAs were verified by sequencing and inserted into the eukaryotic expression vector pKCR2²² as described⁹.

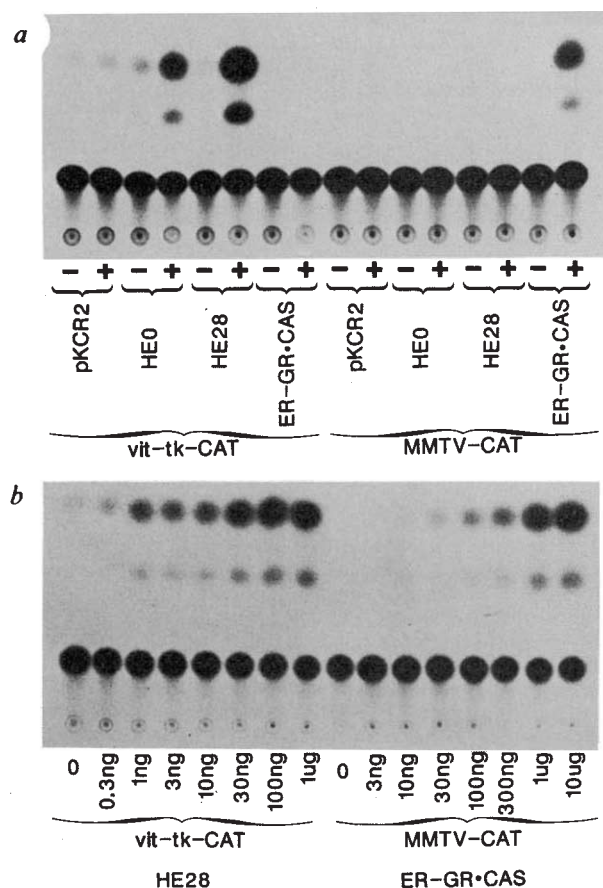


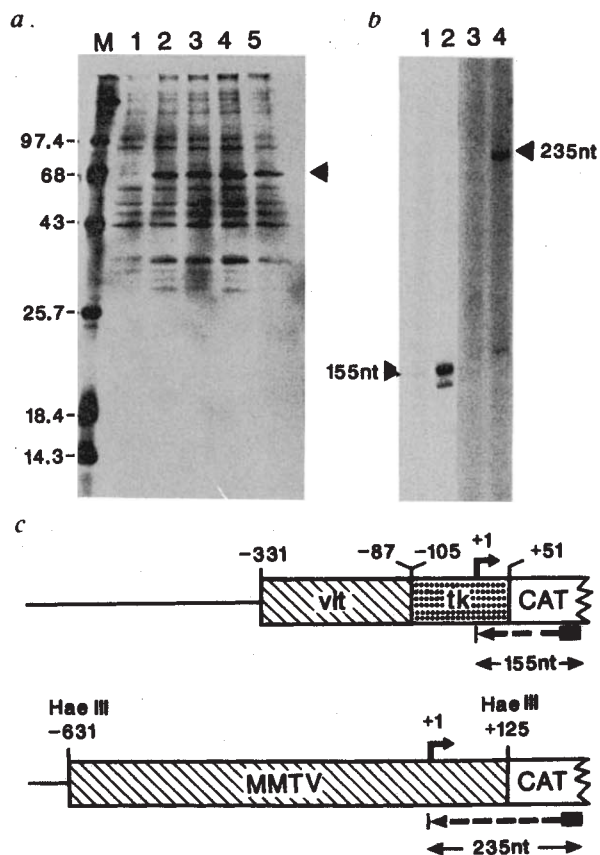
Fig. 2 **a**, The receptor chimaera, ER-GR.CAS, activates the glucocorticoid-responsive reporter gene (MMTV-CAT) in the presence of oestradiol. Each of the expression vectors (pKCR2, HEO, HE28 and ER-GR.CAS) were co-transfected into HeLa cells together with either the oestrogen-responsive (*vit-tk-CAT*) or the glucocorticoid-responsive (MMTV-CAT) reporter genes in the absence (-) or presence (+) of 10^{-8} M oestradiol. **b**, Titration of the amount of the receptor expression vector required to stimulate CAT activity. The indicated amounts of plasmid DNA of the hER mutant (HE28) or the receptor chimaera (ER-GR.CAS) were co-transfected into HeLa cells together with either *vit-tk-CAT* or MMTV-CAT, respectively, in the presence of 10^{-8} M oestradiol. **Methods.** Experiments to demonstrate the activation specificity of the receptors HEO, HE28 and ER-GR.CAS were performed as follows. HeLa cells (10^6 cells) were co-transfected³ with 1 μ g of pKCR2, HEO, HE28 or ER-GR.CAS, together with 1 μ g of the indicated reporter gene with the total amount of DNA adjusted to 20 μ g with Bluescribe M13+. After 24 h the calcium phosphate-DNA precipitate was removed and fresh medium, with or without 10^{-8} M oestradiol, added. After a further 24 h the cells were collected and cell extracts prepared by three cycles of freeze-thawing followed by centrifugation at 10,000g for 5 min. CAT assays were performed for 1 h as described²³ using 25 μ g of protein extract and 0.1 μ Ci of [¹⁴C]chloramphenicol (Amersham). The titration experiments were performed as described above except that the CAT assays were performed using 10 μ g of protein extract and 0.05 μ Ci of [¹⁴C]chloramphenicol.

recombinant (refs 10, 16 and Fig. 2a), but has no effect upon the glucocorticoid-responsive element (GRE) of the mouse mammary tumour virus (MMTV) LTR present in the MMTV-CAT reporter gene (Figs. 2a and 3c). To facilitate the exchange of region C between the two receptors, unique restriction enzyme recognition sites were created at the 5' (*KpnI*) and 3' (*XhoI*) ends of this region in the complementary DNAs of both the hER and hGR (see Fig. 1a, HE28 and GR.CAS). This allows

the DNA encoding the cysteine-rich region to be exchanged as a 'cysteine cassette' (CAS) without disturbing the reading frame. The insertions of the *KpnI* and *XhoI* sites introduce respectively two (Gly and Thr) and three (Ala, Arg and Glu) new amino acids. After transfection of HE28 into HeLa cells, the mutant hER protein was apparently expressed as efficiently as the wild-type HEO receptor (Fig. 3a, lanes 2 and 3) and bound oestradiol (data not shown). Furthermore, co-transfection of

Fig. 3 a, Fluorograph showing the size of the transiently expressed receptors in HeLa cells. Lanes 1–5 correspond to transfection with pKCR2, HEO, HE28, ER-GR.CAS and HE27, respectively. The position of the intact receptor (~66k) is shown by an arrowhead at the right hand side. Lane M represents the markers (phosphorylase B, 97.4k; BSA, 68k; ovalbumin, 43k; α -chymotrypsinogen, 25.7k; β -lactoglobulin, 18.4k; lysozyme, 14.3kDa). **b**, Primer extension analysis of the oestradiol-induced CAT transcripts. HE28 (lanes 1 and 2) or ER-GR.CAS (lanes 3 and 4) were co-transfected into HeLa cells together with *vit-tk*-CAT or MMTV-CAT, respectively, in the absence (lanes 1 and 3) or presence (lanes 2 and 4) of 10^{-8} M oestradiol. Arrowheads, position of the primer extension product of *vit-tk*-CAT (155 nucleotides (nt)) and MMTV-CAT (235 nt) when transcription is initiated at the correct start site. **c**, Structure of the hormone-responsive reporter genes. The organisation of the oestradiol-responsive (*vit-tk*-CAT)¹⁵ and glucocorticoid responsive (MMTV-CAT)²⁰ reporter genes, is shown schematically. The plasmid *vit-tk*-CAT is composed of the upstream sequence of the *Xenopus vitellogenin A2* gene (–331 to –87) (*vit*), the Herpes simplex virus thymidine kinase promoter (–105 to +51) (*tk*) and the bacterial chloramphenicol acetyl transferase (CAT) coding sequence. MMTV-CAT consists of the mouse mammary tumour virus LTR (–631 to +125) (MMTV) fused upstream of the CAT gene. The start site of each gene (+1) is marked with an arrow. The oligonucleotide complementary to the CAT sequence (see below), which was used for primer extension is shown (small solid block) together with the primer extension product (dashed line).

Methods. **a**, To determine the size of the expressed receptors, HeLa cells (10^6 cells) were transfected with 5 μ g of each of the receptor expression vectors together with 15 μ g of Bluescribe M13+. The cells were processed as described previously³ except that 1 μ g each of the monoclonal antibodies H222 and H226 were used during the immunoabsorption of the receptor proteins. **b**, For the primer extension analysis, HeLa cells (10^6 cells) were co-transfected with either 1 μ g of HE28 and 5 μ g of *vit-tk*-CAT or 15 μ g of ER-GR.CAS together with 5 μ g of MMTV-CAT. After 24 h, fresh medium, with or without 10^{-8} M oestradiol, was added. Cytoplasmic RNA²⁴ was isolated 24 h later and 50 μ g used as the template for primer extension²⁵ using the oligonucleotide 5'-GTTCTTTAC-GATGCCATTGGGATATATCAACGG-3'.



the HE28 expression vector together with the *vit-tk*-CAT reporter gene into HeLa cells resulted in oestradiol-dependent stimulation of CAT activity similar to that obtained using HEO (Fig. 2a, and data not shown). No effect upon CAT activity was observed using the parent expression vector, pKCR2, which does not contain hER sequences (Fig. 2a). HE28, like HEO, was unable to activate the glucocorticoid-responsive MMTV-CAT reporter gene (Fig. 2a).

We next exchanged the cysteine cassette of the hER with that of the hGR, producing the ER-GR.CAS chimaera (Fig. 1a). Transient expression of ER-GR.CAS in HeLa cells resulted in the synthesis of a protein of the expected size (Fig. 3a, lane 4), which bound oestradiol (data not shown), but could not activate the oestradiol-responsive reporter gene *vit-tk*-CAT, in the absence or presence of oestradiol (Fig. 2a). In contrast, ER-GR.CAS did activate the glucocorticoid-responsive reporter gene MMTV-CAT and oestradiol was required for this activation (Fig. 2a). Under the conditions described, 100 ng of HE28 fully activated the *vit-tk*-CAT reporter gene, whereas at least 1 μ g of ER-GR.CAS was required to reach full activation of MMTV-CAT (Fig. 2b). This difference in efficiency may reflect a difference in the affinity of each receptor for its respective responsive element and/or between the amount of the two receptors synthesized (although similar amounts were produced when transfecting 5 μ g of receptor DNA; Fig. 3a, lanes 3 and 4). In co-transfections of HEO with *vit-tk*-CAT, half-maximal CAT activity was obtained at ~100 pM oestradiol¹⁰, a value very similar to that required for half-maximal induction of the oestrogen-responsive pS2 gene in the human breast cancer cell line MCF-7¹⁷. A similar concentration of oestradiol was required in the case of HE28 and ER-GR.CAS to obtain half-maximal

expression of *vit-tk*-CAT and MMTV-CAT, respectively (data not shown), suggesting that oestradiol acts by a similar mechanism on both the hER and the ER-GR.CAS chimaeric receptor. Primer extension experiments were used (Fig. 3b,c) to verify that transcription was initiated at the correct start site of the *vit-tk*-CAT (Fig. 3b, lane 2) and MMTV-CAT recombinants (Fig. 3b, lane 4), using the HE28 and ER-GR.CAS receptors, respectively. These experiments also confirmed the oestradiol-dependence of the ER-GR.CAS-induced MMTV-CAT transcription (compare lanes 3 and 4).

We conclude from this study that region C of steroid receptors determines the specificity of target gene activation by the hormone. Indeed, it is likely that region C of any given receptor interacts directly with its cognate hormone-responsive element. It has been shown that the glucocorticoid and progesterone receptors, which are by far the most homologous within region C (91%; ref. 7), recognize common DNA-binding sites^{18,19} and activate a common hormone-responsive element²⁰. As the hER does not activate the glucocorticoid-responsive MMTV-CAT reporter gene (see above), we assume that the amino acids which are conserved in region C of all receptors⁷ provide the structural scaffolding of this domain, whereas the variable amino acids determine the specificity of the receptor-promoter element interaction. In this respect, it will be interesting to know whether the two halves of region C (see Fig. 1b), which are encoded in separate exons conserved in both the oestrogen (M. Ponglikitmongkol, S.G. and P.C., unpublished results) and progesterone⁷ receptors, are both required for DNA-binding and in determining the specificity of receptor-DNA interaction. Further studies involving receptors containing either of these halves or a combination of ER with GR halves are in progress. Further studies

are also required to investigate the significance of the inactivation of the ER when the second pair of cysteines are replaced with histidines in the left-hand domain of region C. Should this be taken as an indication that zinc-stabilized 'DNA-binding fingers' are not present in steroid hormone receptors, or as a reflection of deleterious changes in the finger structure after such substitutions?

Exchange of the hER region C with that of the hGR results in oestradiol-dependent activation of gene expression. Thus we also conclude that, whatever is the mechanism underlying activation of transcription by the hormone-receptor complex, it allows the oestradiol-binding domain of the hER to co-operate efficiently with the putative DNA-binding domain of the hGR to produce an oestradiol-regulatable receptor. If interactions between the two domains are essential for regulation of receptor activity by the hormone, then our results indicate that these interactions have been conserved during evolution. Note, however, that these results have been obtained with artificial hormone-responsive reporter genes present as plasmids in the transfected cells. Further studies on the induction of glucocorticoid-responsive cellular genes, in their natural chromatin environment, by the ER-GR-CAS receptor in the presence of oestradiol, are required to establish whether this chimera completely mimics the effect of the natural glucocorticoid receptor.

One important consequence of the results reported here is the possibility of creating chimaeras between *trans*-acting transcriptional regulatory factors that contain sequences homologous to the steroid hormone receptor region C. Expressing such novel transcriptional regulatory proteins in cultured cells and transgenic animals provides a new approach for studying the control of gene expression during embryonic development and terminal differentiation.

We thank M. Govindan, M. Devic and U. Stropp for providing the hGR cDNA fragment, G. Ryffel and H. Ponta for providing *vit-tk*-CAT and MMTV-CAT, respectively, and Abbott Laboratories for providing the H222 and H226 monoclonal antibodies. We also thank V. Kumar for helpful discussion, I. Issemann and E. Scheer for technical assistance, A. Staub for synthesizing the oligonucleotides, the tissue culture lab for media and cells, and C. Werlé and B. Boulay for the preparation of the figures. This research was supported by INSERM, CNRS (ATP Etats-Unis) and the Ministère de la Recherche et de la Technologie. S.G. was a recipient of a Royal Society European Exchange fellowship during part of this work.

Received 7 November; accepted 21 November 1986.

1. Yamamoto, K. R. *A. Rev. Genet.* **19**, 209-252 (1985).
2. Beato, M. *et al. Chromosomal Proteins and Gene Expression* (eds Reece, G. R., Goodwin, G. H. & Puigdomenech, P.) 121-143 (Plenum, New York, 1985).
3. Green, S. *et al. Nature* **320**, 134-139 (1986).
4. Krust, A. *et al. EMBO J.* **5**, 891-897 (1986).
5. Hollenberg, S. M. *et al. Nature* **318**, 635-641 (1985).
6. Miesfeld, R. *et al. Cell* **46**, 389-399 (1986).
7. Jeltsch, J. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5424-5428 (1986).
8. Conneely, O. M. *et al. Science* **233**, 767-770 (1986).
9. Kumar, V., Green, S., Staub, A. & Chambon, P. *EMBO J.* **5**, 2231-2236 (1986).
10. Kumar, V. *et al.* (manuscript in preparation).
11. Giguère, V., Hollenberg, S. M., Rosenfeld, M. G. & Evans, R. M. *Cell* **46**, 645-652 (1986).
12. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609-1614 (1985).
13. Hartshorne, T. A., Blumberg, H. & Young, E. T. *Nature* **320**, 282-287 (1986).
14. Weinberger, C., Hollenberg, S. M., Rosenfeld, M. G. & Evans, R. M. *Nature* **318**, 670-672 (1985).
15. Klein-Hitpass, L., Schorpp, M., Wagner, U. & Ryffel, G. U. *Cell* **46**, 1053-1061 (1986).
16. Druge, P. M. *et al. Nucleic Acids Res.* (in the press).
17. Westley, B. *et al. J. biol. Chem.* **259**, 10030-10035 (1984).
18. Von der Ahe, *et al. Nature* **313**, 706-709 (1985).
19. Von der Ahe, D., Renoir, J. M., Buchou, T., Baulieu, E. E. & Beato, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2817-2821 (1986).
20. Cato, A. C. B., Miksick, R., Schütz, G., Arnemann, J. & Beato, M. *EMBO J.* **5**, 2237-2240 (1986).
21. Govindan, M. V., Devic, M., Green, S., Gronemeyer, H. & Chambon, P. *Nucleic Acids Res.* **13**, 8293-8304 (1985).
22. Breathnach, R. & Harris, B. A. *Nucleic Acids Res.* **11**, 7119-7136 (1983).
23. Gorman, C. M., Merlino, G. T., Willingham, M. C., Pastan, I. & Howard, B. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6777-6781 (1982).
24. Groudine, M., Peretz, M. & Weintraub, H. *Molec. cell Biol.* **1**, 281-288 (1981).
25. Nelson, C. *et al. Nature* **322**, 557-562 (1986).

Blebbing, free Ca^{2+} and mitochondrial membrane potential preceding cell death in hepatocytes

John J. Lemasters, James DiGuiseppi,
Anna-Liisa Nieminen & Brian Herman

Laboratories for Cell Biology, Department of Anatomy,
School of Medicine, University of North Carolina,
Chapel Hill, North Carolina 27514, USA

Cell surface 'blebbing' is an early consequence of hypoxic and toxic injury to cells¹⁻⁵. A rise in cytosolic free Ca^{2+} has been suggested as the stimulus for bleb formation³⁻⁵ and the final common pathway to irreversible cell injury⁶⁻⁹. Here, using digitized low-light video microscopy, we examine blebbing, cytosolic free Ca^{2+} , mitochondrial membrane potential and loss of cell viability in individual cultured hepatocytes. Unexpectedly, we found that after 'chemical hypoxia' with cyanide and iodoacetate, cytosolic free Ca^{2+} does not change during bleb formation or before loss of cellular viability. Cell death was precipitated by a sudden breakdown of the plasma membrane permeability barrier, possibly caused by rupture of a cell surface bleb.

Hepatocytes isolated from rat liver by collagenase perfusion¹⁰ were cultured overnight, loaded with fura-2-acetoxymethyl ester and rhodamine 123, and transferred to propidium iodide-containing medium. Endogenous cytosolic esterases release free fura-2 which remains trapped in the cytosolic compartment. Fura-2 displays a Ca^{2+} -dependent fluorescence excited at 350 nm and a Ca^{2+} -independent fluorescence excited at 365 nm^{11,12}. We used the ratio of fluorescence at these two wavelengths to monitor cytosolic free Ca^{2+} quantitatively. The ratio method corrects for variable loading of probe and differential volume fractions in the two-dimensional image. The two other fluorescence probes used were rhodamine 123, a probe of mitochondrial membrane potential¹³, and propidium iodide which selectively stains the nuclei of nonviable cells¹⁴.

The fluorescence of these probes was measured with a digitized low-light video microscope system consisting of an inverted phase/fluorescence microscope, a low-light level camera and video digitizing equipment for frame averaging, background subtraction, ratio imaging and data storage¹⁵. This system allowed us to work at very low levels of excitation energy and fluorescent probe concentration. At the probe concentration used photobleaching, photodamage and disruption of cellular activities were essentially eliminated.

In hypoxia, ATP production by aerobic metabolism is inhibited and cellular ATP levels fall rapidly. To duplicate this condition, cyanide and iodoacetate were added to cultured hepatocytes to inhibit ATP formation by oxidative phosphorylation and glycolysis. Phase and fluorescent images were collected from fields of several cells at frequent intervals; each cell was scored for surface blebbing from phase images. Ratio images of fura-2 fluorescence were converted to Ca^{2+} concentrations and averaged for individual cells. Similarly, rhodamine 123 fluorescence was quantified. Finally, individual cells were scored for propidium iodide uptake. Only cells which initially excluded propidium iodide were analysed.

Within minutes after addition of cyanide and iodoacetate, clear homogenous blebs are evident in phase images of the cells (Fig. 1). By 30 minutes, more than 70% of originally viable cells show blebbing. During bleb development, rhodamine 123 fluorescence declines steadily, but cytosolic free Ca^{2+} is unchanged and only a small increase of propidium iodide uptake occurs.

After 30 minutes, the percentage of cells exhibiting blebs begins to decrease and the proportion of nonviable cells to increase. By 40 minutes, most cells have lost viability and surface

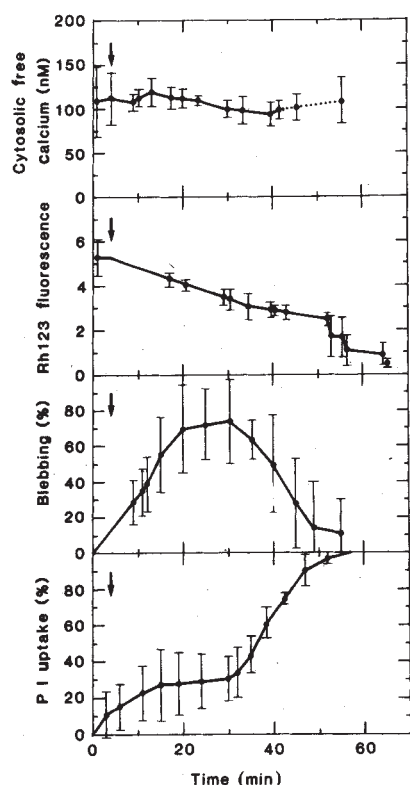


Fig. 1 Changes in cytosolic free Ca^{2+} , rhodamine 123 fluorescence, cell surface blebbing, and propidium iodide uptake after chemical hypoxia in cultured hepatocytes. Hepatocytes were labelled with fura-2 and rhodamine 123 and incubated in propidium iodide (PI)-containing medium. Arrow, addition of 5 mM KCN and 10 mM iodoacetate to initiate 'chemical hypoxia'. Dotted line (upper panel), first loss of fura-2 in individual experiments. At the termination of the line, fura-2 was lost from cells in all experiments. Rhodamine 123 fluorescence, which indicates mitochondrial membrane potential, decreases at an almost constant rate after addition of KCN and iodoacetate. After cell death, virtually all remaining rhodamine 123 escaped from the cells.

Methods. Hepatocytes were isolated from the livers of fed, 200–300 g male Sprague-Dawley rats and cultured overnight at 37 °C with 5% CO_2 on collagen-coated glass coverslips in Waymouth media MB-752/2 containing 100 nM insulin and 1% fetal calf serum. Cultures were incubated for 30 min with fura-2-acetoxymethyl ester (3.3 μM) and rhodamine 123 (790 nM). Cytosolic fura-2 concentration with this loading protocol was 185 μM as determined in hepatocyte suspensions by digitonin-releasable, Ca^{2+} -responsive absorbance at the wavelength pair 340–380 nm. After three washes with Hanks medium, the coverslips were transferred to Krebs-Henseleit-bicarbonate medium buffered with 20 mM NaHEPES, pH 7.4, containing 50 nM propidium iodide. The coverslips were viewed with a Zeiss IM35 microscope using a 63 $\times$ N.A. 1.3 objective. After collecting baseline data, iodoacetate and KCN (neutralized) were added. Excitation light for fluorescence was provided by a 100-watt Hg lamp, and temperature was maintained at 37 °C with an air curtain incubator. Video images were obtained by averaging up to 256 frames with the following filter combinations: fura-2—excitation, 350 nm or 365 nm, emission, >450 nm; rhodamine 123—excitation, 490 nm, emission, >530 nm; propidium iodide—excitation, 554 nm, emission, >580 nm. Video frames were collected using an ISIT-66 camera (DAGE-MTI, Inc., Michigan City, IN) and digitized with an IP-512 series board set (Imaging Technologies Inc.) housed in an 11/23 based Q-Bus microcomputer (Scientific Micro System, Inc.). A technical description of our instrumentation is given elsewhere¹⁵. Routinely, the excitation intensity was attenuated 100–1,000 fold before reaching the cells, and background images (areas on the coverslip devoid of cells or an empty chamber containing medium from a treated cell population) were obtained at the beginning and end of each experiment. Calcium concentration was measured by subtracting background from images taken at 350 and 365 nm and then dividing the resultant images on a pixel by pixel basis. To obtain cytosolic free Ca^{2+} concentrations for an individual cell, the mean value of the pixel ratios for the cell was compared with values obtained with the same equipment using fura-2-containing EGTA- Ca^{2+} buffers. The 350/365 ratio increased hyperbolically fivefold over the range 0–500 nM free Ca^{2+} with a half-maximal response at 225 nM. In some experiments, 380 nm fluorescence decreased to background levels as cytosolic free Ca^{2+} increased after hormonal stimulation (data not shown). This indicates that virtually all intracellular fura-2 was accessible to Ca^{2+} . Cells taking up propidium iodide before KCN and iodoacetate addition were not counted. Results shown are running averages ($\pm$ s.d.) from experiments with four different preparations of hepatocytes. Three to six cells were counted per experiment.

blebs have almost disappeared. As viability is lost, two of the probe molecules leak from the cells—first fura-2 (dotted line in Fig. 1) and then rhodamine 123 shortly afterwards. In control incubations blebs were absent, propidium iodide uptake did not increase, and cytosolic free Ca^{2+} and rhodamine 123 fluorescence were stable for over an hour. Both cyanide and iodoacetate are required for these changes. Cyanide alone caused no change in cell morphology, viability or cytosolic free Ca^{2+} after incubations of over an hour.

The sequence of events is illustrated for an individual cell in Fig. 2. Before addition of cyanide and iodoacetate, viable cells have dense granular cytoplasm, bright fura-2 and rhodamine 123 fluorescence, and exclude propidium iodide (Fig. 2a–d). Fura-2 fluorescence is diffuse, consistent with its cytosolic location, whereas rhodamine 123 fluorescence is granular, labelling individual mitochondria. After several minutes of chemical hypoxia with cyanide and iodoacetate, numerous small blebs containing Ca^{2+} -specific fura-2 fluorescence but not rhodamine 123 fluorescence develop on the surfaces of viable cells (Fig. 2e–g). As blebs form, rhodamine 123 fluorescence decreases throughout the cytoplasm but Ca^{2+} -specific fluorescence is unchanged. In most cells, one or a few large terminal blebs form as the injury progresses further (Fig. 2h). Subsequently the terminal blebs rupture or are reabsorbed, then fura-2 leaks out of the cytoplasm and propidium iodide is taken up into the nucleus (Fig. 2i–n). These events indicate an abrupt transition from reversible to irreversible injury.

No change of cytosolic free Ca^{2+} precedes bleb formation or the transition to irreversible injury. Several positive control experiments were undertaken to ensure that ratio images of fura-2 fluorescence are indeed responsive to free Ca^{2+} . As expected, cytosolic free Ca^{2+} calculated from fura-2 fluorescence ratios increases after stimulation with the Ca^{2+} -mobilizing hormones, phenylephrine, vasopressin and epidermal growth factor (Table 1)^{16–20}. Low concentrations of ionomycin, a Ca^{2+} ionophore, also increase cytosolic free Ca^{2+} , whereas EGTA, a Ca^{2+} chelator, decreases Ca^{2+} (Table 1). Thus, fura-2 was responsive to both upward and downward changes of cytosolic free Ca^{2+} . Note that despite changes in free Ca^{2+} caused by these treatments, blebbing does not occur and viability is not lost.

In non-liver systems, the calcium-sensitive photoprotein aequorin has been used to measure cytosolic free Ca^{2+} during anoxia and 'chemical hypoxia'. In isolated ventricle myocytes microinjected with aequorin, rounding-up and blebbing of the surface occurs after addition of cyanide and 2-deoxyglucose, but aequorin luminescence remains at resting levels indicating no change in cytosolic free Ca^{2+} (refs 20, 21). Similar observations have been made in papillary muscle^{22,23}. These findings have been questioned on the basis that the detection threshold of aequorin-microinjected cells may be too high (>160 nM) to see a rise of free Ca^{2+} (refs 24, 25). In perfused monkey kidney epithelial cells loaded with aequorin by hypo-osmotic shock^{24,25}, luminescence increases steadily during anoxia, indicating an

Table 1 Maximum changes in cytosolic free Ca^{2+} after addition of various agents

Addition	Change of cytosolic free Ca^{2+} (%)
5 μM phenylephrine	90 $\pm$ 13% (5)
100 nM vasopressin	244 $\pm$ 27% (2)
10 nM epidermal growth factor	90 $\pm$ 10% (4)
1.5 μM ionomycin	121 $\pm$ 41% (4)
5 mM EGTA	–22 $\pm$ 11% (6)

Cultured hepatocytes were loaded with fura-2, and cytosolic free Ca^{2+} was determined for individual cells from ratio images of fura-2 fluorescence as described (Fig. 1). Ratio images were recorded from about 10 s to about 10 min after additions. Values are maximum percent changes $\pm$ s.d. from the number of experiments shown in parentheses. Initial cytosolic free Ca^{2+} was 112 nM $\pm$ 6 nM (s.e.m.).

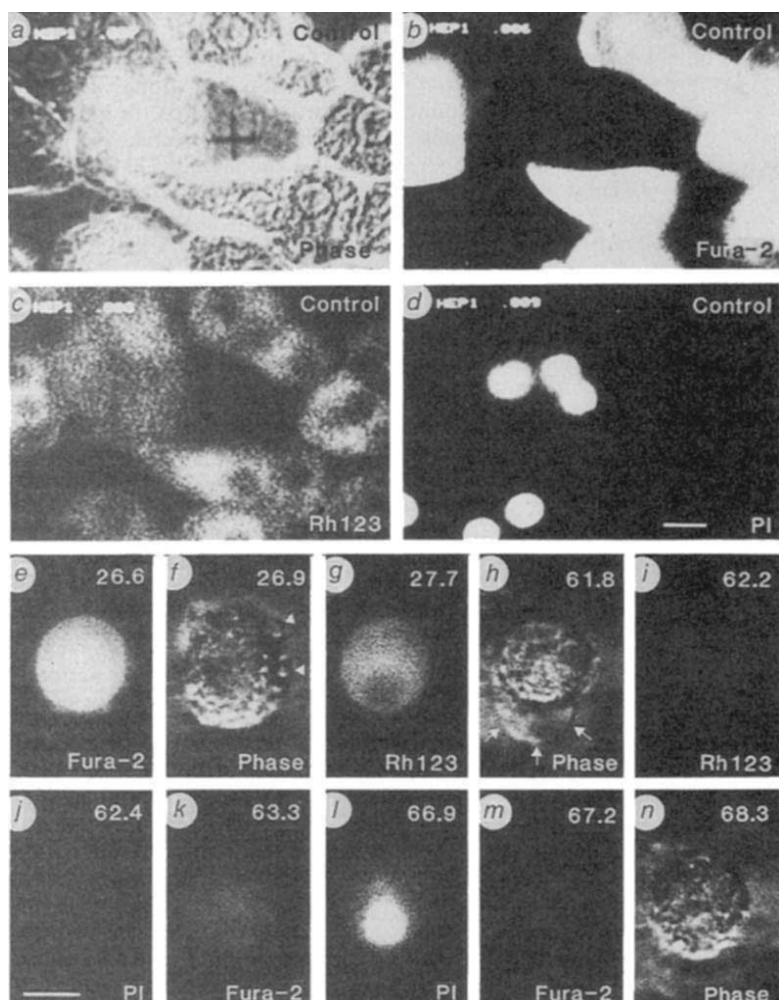


Fig. 2 Video micrographs of cultured hepatocytes during chemical hypoxia. Cultured hepatocytes were incubated as described (Fig. 1). In the absence of iodoacetate and KCN (a–d), viable cells were characterized by granular cytoplasm in the phase image (a), bright fura-2 fluorescence at the Ca^{2+} -specific excitation wavelength of 350 nm (b), bright mitochondria-specific rhodamine 123 fluorescence (c) and exclusion of propidium iodide (d). Nonviable cells were pale by phase microscopy, poorly fluorescent for fura-2, and brightly stained with nuclear propidium iodide. The nonviable cells shown here retain some rhodamine 123 fluorescence suggesting that loss of viability was quite recent, possibly during manipulations to mount the cells on the microscope. In e–n, the progression of cell injury and onset of cell death are documented for a single hepatocyte following exposure to iodoacetate and cyanide. At about 27 minutes, blebs were clearly evident on the cell surface (f, arrowheads). Fura-2 fluorescence but not rhodamine 123 fluorescence extended into the blebs (e, g, and data not shown; see also refs 2, 4). In this cell, loss of viability occurred after about 65 min. Before loss of viability, a single large, clear bleb was observed on the cell surface (h, arrows). Rhodamine 123 fluorescence (i) was diminished but cell death had not occurred as evidenced by an absence of propidium iodide uptake (j). Near the time of cell death, fura-2 fluorescence (k) decreased markedly, and the nucleus became labelled with propidium iodide (l). After cell death, fura-2 fluorescence and the large terminal bleb observed 6.5 minutes earlier were absent (m, n). Numbers in the upper right corners of panels e–n indicate minutes following addition of iodoacetate and KCN. Control frames (a–d) were obtained before addition of inhibitors. Bars in d and j are 15 μm .

increasing cytosolic free Ca^{2+} . But the question of whether this change was a small increase in all the cells or a large increase in a small hyperpermeable subpopulation remains. The present study avoids these problems by examining cells individually with increased sensitivity. Decreases as well as increases in cytosolic Ca^{2+} concentrations are detectable.

Phosphorylase *a* is activated in suspensions of isolated hepatocytes in parallel with bleb formation after oxidative stress with various cytotoxic chemicals^{5,9}. Because phosphorylase *a* activation is a Ca^{2+} -dependent process, it was concluded that an increase of cytosolic free Ca^{2+} initiates formation of cell surface blebs. Other factors influence phosphorylase activation, however, and comparison between cytosolic free Ca^{2+} measured by quin-2 and phosphorylase activity in isolated hepatocytes does not show a close correlation¹⁷. Our results also show that an increase of cytosolic free Ca^{2+} is not a requirement for cell surface blebs to form.

Here, for the first time, we relate the onset of irreversible injury to the time course of changes in cell surface morphology, cytosolic free Ca^{2+} and mitochondrial membrane potential. Our model of 'chemical hypoxia' to cultured hepatocytes compares closely to hypoxic injury to intact, perfused livers^{2,4,26}. In both systems, blebs appear within minutes of metabolic insult, but irreversible damage does not occur for about 30 minutes. The requirement of both cyanide and iodoacetate for loss of viability indicates that metabolism of endogenous glycogen is capable of maintaining the viability of cultured hepatocytes. Similarly, in intact glycogen-rich livers from fed rats hypoxic perfusions of up to 120 minutes fail to cause hepatocellular death, whereas identical perfusions of glycogen-depleted livers from starved rats cause substantial cell death after as little as 30 minutes²⁶.

In our hands, the only change accompanying bleb formation is a fall of the mitochondrial membrane potential. In the absence of inhibition of the mitochondrial ATPase, the mitochondrial membrane potential should, more or less, be proportional to ΔG_p , the phosphorylation potential, which is the free energy change of ATP synthesis. Polymerization of microfilaments and microtubules requires high energy nucleoside triphosphates, so cytoskeletal disruption during bleb formation may simply be a consequence of a deficiency of ATP and other high energy compounds. This is plausible because cellular microfilaments and microtubules are in a dynamic state of continuous formation and disassembly²⁷.

Before cell death, rhodamine 123 fluorescence within cells decreases by about 50% (Fig. 1) and diffuse rhodamine 123 fluorescence outside cells increases by about 60% (data not shown). Because rhodamine 123 distributes electrophoretically across the mitochondrial membrane as a membrane-permeable cation²⁸, by the Nernst equation these changes correspond to a decrease of membrane potential of about 30 mV or 20% assuming an initial value of 160 mV^{29,30}. Such a small decrease seems implausible; however, a commensurate decrease of cytosolic ΔG_p corresponds to a 99% fall of ATP/ADP³¹. Thus, a relative persistence of the mitochondrial membrane potential should be expected even in the face of precipitously falling ATP levels. Andersson and Jones³⁰ demonstrated persistence of the mitochondrial membrane potential during anoxia in suspensions of isolated hepatocytes using different methods. This persisting mitochondrial membrane potential may be important for Ca^{2+} homeostasis by hypoxic cells.

The progression from reversible to irreversible injury was associated with and possibly caused by rupture of terminal blebs

as blebs are no longer evident in phase images following the loss of viability. On this basis, a minimal hypothesis for the progression of cell injury and death during hypoxia can be constructed: (1) hypoxia, (2) inhibition of ATP synthesis, (3) fall of ΔG_p , (4) disturbance of membrane-cytoskeleton connections, (5) focal weakening of the cell surface (small blebs), (6) expansion and/or coalescence of weakened areas (large terminal blebs), (7) rupture of terminal blebs leading to membrane hyperpermeability, and (8) collapse of ionic gradients, irreversible

loss of metabolic intermediates and cytosolic enzymes, and damage of mitochondria (cell death).

We thank Dr H. Shelton Earp for hepatocytes, Dr Lerson Tanasugarn for helpful discussions and Dr Roger Madison for bringing propidium iodide to our attention. This work was supported in part by Grants AM37034 and HL35490 from the National Institutes of Health, Grant 86-1299 from the American Heart Association and the Whitaker Foundation. J.J.L. is an Established Investigator of the American Heart Association.

Received 28 April; accepted 20 October 1986.

1. Trump, B. F., Penttilä, A. & Berezsky, I. K. *Virchows Arch. B Cell Path.* **29**, 281-307 (1979).
2. Lemasters, J. J., Ji, S. & Thurman, R. G. *Science* **213**, 661-663 (1981).
3. Jewell, S. A., Bellomo, G., Thor, H. & Orrenius, S. *Science* **217**, 1257-1259 (1982).
4. Lemasters, J. J., Stemkowski, C. J., Ji, S. & Thurman, R. G. *J. Cell Biol.* **97**, 778-786 (1983).
5. Thor, H., Hartzell, P. & Orrenius, S. *J. Biol. Chem.* **259**, 6612-6615 (1985).
6. Schanne, F. A. X., Kane, A. B., Young, E. E. & Farber, J. L. *Science* **206**, 700-702 (1979).
7. Trump, B. F. *et al.* in *Scanning Electron Microscopy* (ed. A. M. F. O'Hare) 437-462 (SEM Inc., Chicago, 1980).
8. Farber, J. L. *Lab. Invest.* **47**, 114-123 (1982).
9. Bellomo, G. & Orrenius, S. *Hepatology* **5**, 876-882 (1985).
10. Hochachka, P. W. *Science* **231**, 234-241 (1986).
11. Gryniewicz, G., Poenie, M. & Tsien, R. Y. *J. Biol. Chem.* **260**, 3440-3450 (1985).
12. Williams, D. A., Fogarty, K. E., Tsien, R. Y. & Fay, F. S. *Nature* **318**, 558-561 (1985).
13. Johnson, L. V., Walsh, M. L., Bockus, B. J. & Chen, L. B. *J. Cell Biol.* **88**, 526-535 (1981).
14. Krishan, A. *J. Cell Biol.* **66**, 188-193 (1975).
15. DiGiuseppi, J., Inman, R., Ishihara, A., Jacobson, K. & Herman, B. *BioTechniques* **3**, 394-403 (1985).

16. Charest, R., Blackmore, P. F., Berthon, B. & Exton, J. H. *J. Biol. Chem.* **258**, 8769-8773 (1983).
17. Binet, A., Bertho, B. & Claret, M. *Biochem. J.* **228**, 565-574 (1985).
18. Morris, J. D. H., Metcalfe, J. C., Smith, G. A., Hesketh, T. R. & Taylor, M. V. *FEBS Lett.* **169**, 189-193 (1984).
19. Moolenaar, W. H., Aerts, R. J., Tertoolen, L. G. J. & de Laat, S. W. *J. Biol. Chem.* **261**, 279-284 (1986).
20. Cobbold, P. H. & Bourne, P. K. *Nature* **312**, 444-446 (1984).
21. Cobbold, P. H., Bourne, P. K. & Cuthbertson, K. S. R. *Basic Res. Cardiol.* **80**, (Suppl. 2), 155-158 (1985).
22. Allen, D. G. & Orchard, C. H. *J. Physiol.* **339**, 107-122 (1983).
23. Allen, D. G. & Orchard, C. H. *J. molec. Cell Cardiol.* **16**, 117-128 (1984).
24. Borle, A. B. & Snowdowne, K. W. *Science* **217**, 252-254 (1982).
25. Snowdowne, K. W., Freudenrich, C. C. & Borle, A. B. *J. Biol. Chem.* **260**, 11619-11626 (1985).
26. Bradford, B. U., Marotto, M., Lemasters, J. J. & Thurman, R. G. *J. Pharmac. exp. Ther.* **236**, 263-268 (1986).
27. Alberts, B. *et al. Molecular Biology of the Cell* (Garland, New York, 1983).
28. Emaus, R. K., Grunwald, R. & Lemasters, J. J. *Biochim. biophys. Acta* **850**, 436-448 (1986).
29. Hoek, J. B., Nicholls, D. G. & Williamson, J. R. *J. Biol. Chem.* **255**, 1458-1464 (1980).
30. Andersson, B. S. & Jones, D. P. *Analyt. Biochem.* **146**, 164-172 (1985).
31. Veech, R. L., Lawson, J. W. R., Cornell, N. W. & Krebs, H. A. *J. Biol. Chem.* **254**, 6538-6547 (1979).

A transcript from a *Drosophila* pattern gene predicts a protein homologous to the transforming growth factor- β family

Richard W. Padgett, R. Daniel St. Johnston & William M. Gelbart

Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

The decapentaplegic gene complex (DPP-C) has been implicated in several events in pattern formation during *Drosophila* development¹⁻³. During embryogenesis, the DPP-C participates in the establishment of dorsal-ventral specification. Later, it is required for the correct morphogenesis of the imaginal disks, which will form much of the adult epidermis. We have undertaken a molecular analysis of the DPP-C to determine what role it plays in positional information. It appears to be a large genetic unit (>40 kilobases (kb)) consisting mostly of cis-regulatory information controlling the expression of a set of overlapping transcripts that differ at their 5' ends, but share the bulk of their transcribed sequences. Here, we describe the sequence analysis of two complementary DNAs comprising 4.0 kb of a 4.5-kb transcript. The C-terminus of the protein thereby deduced exhibits strong sequence homology (25-38% amino-acid identity) to the C-termini of a class of mammalian proteins that includes transforming growth factor- β (TGF- β)^{4,5}, inhibin⁶⁻⁹ and Müllerian inhibiting substance (MIS)^{10,11}. These proteins act on target cells to produce a variety of responses, such as stimulation or inhibition of cell division or differentiation. The homology suggests that the DPP-C protein contributes to correct morphogenesis as a secreted factor involved in the differential regulation of cell growth. This is the first report of a member of the TGF- β gene family in a non-mammalian organism, and indicates that one or more members of this gene family existed before arthropod and vertebrate lineages diverged.

The decapentaplegic gene complex of *Drosophila* has been

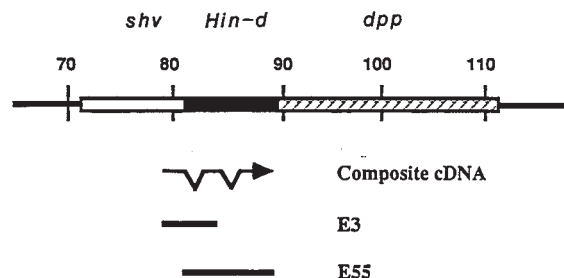


Fig. 1 Map of the DPP-C and location of E3 and E55 cDNAs. The present extent of the DPP-C is shown with the three functional genetic units: shortvein (*shv*); Haplo-insufficiency near decapentaplegic (*Hin-d*) and the decapentaplegic or disk region (*dpp*). The numbers above the DPP-C (in kb) derive from the 'walk' used originally to clone this region. The coding region of the composite cDNA sequence is contained within the *Hin-d* region (solid box) whereas 5' untranslated sequences are located in the *shv* region (open box). The two introns shown in the composite cDNA sequence have been located by comparison with the corresponding genomic DNA sequence (data not shown). The two bars below the composite cDNA correspond to the E3 and E55 cDNAs.

defined on the basis of the phenotypes and allelic interactions of a large array of mutations affecting the development of many pattern elements. The DPP-C contributes to the determination of the dorsal structures in the developing embryo and to the proper morphogenesis of the imaginal disks. The DPP-C is also active in some mesodermal and endodermal derivatives (unpublished observations). Three genetic regions have been identified within this gene complex: shortvein (*shv*), Haplo-insufficiency near decapentaplegic (*Hin-d*) and the decapentaplegic region (*dpp*) (Fig. 1). In homozygotes, *Hin-d* null mutations are early acting embryonic lethal lesions, producing completely ventralized embryos³. Mutations of the decapentaplegic (*dpp*) region result in the loss of adult epidermal structures derived from the centres of the imaginal disks¹. Shortvein mutations engender defects in venation of the wing and in the maxillary palps of the head; many *shv* alleles are also recessive larval lethals².

Fig. 2 Nucleotide sequence and deduced amino-acid sequence of the composite cDNA derived from the E3 and E55 cDNAs. The 3' extent of the E3 cDNA is at position 2,005 and the 5' extent of the E55 cDNA is at position 1,133. After determining the nucleotide sequence of E3 and E55, a single cDNA clone was isolated which matched the composite cDNA presented here. A single ORF 1,778 bp long encodes a protein of 588 amino acids. Three asterisks represent the terminators at the ends of the ORF, also present in the corresponding genomic DNA. An AATAAA potential polyadenylation signal³² is boxed at position 3,917. Single asterisks denote each ten amino acids. The two AUG codons that could initiate translation are boxed at positions 1,188 and 1,587. Three potential proteolytic processing sites, each containing an R-R residue, used in several of the related mammalian genes, are indicated by heavy underlining. □, Potential glycosylation sites (N-X-S or N-X-T)³³. The underlining beginning at position 2,655 indicates the minimum extent of the putative secreted portion of the protein. The amino acid code used is: A, alanine; R, arginine; N, asparagine; D, aspartic acid; C, cysteine; Q, glutamine; E, glutamic acid; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; F, phenylalanine; P, proline; S, serine; T, threonine; W, tryptophan; Y, tyrosine; V, valine.

Methods. The two cDNAs were isolated from a λ gt10 library using standard protocols. The complete sequence of both strands was determined by the dideoxy method using random fragments of the two insert cDNAs cloned into M13 phage derivatives³⁴. Briefly,

The transcriptional pattern of the DPP-C is structurally and temporally complex³. Five transcripts in the size range of 3.5–5.0 kb have been identified with all transcripts overlapping extensively within the two 3′-most exons; both these exons fall within the *Hin-d* region. Four transcripts are initiated within the shortvein region; all transcripts terminate before the *dpp*

region. Two overlapping cDNA clones, E3 and E55, which were isolated from a cDNA library made from poly(A)⁺ RNA extracted from 3–12 hour embryos¹² were sequenced (Fig. 2). Based on Northern blots probed with genomic fragments of the *shv* and *Hin-d* regions (data not shown), the splicing pattern of these cDNAs corresponds to the 4.5-kb DPP-C transcript³,

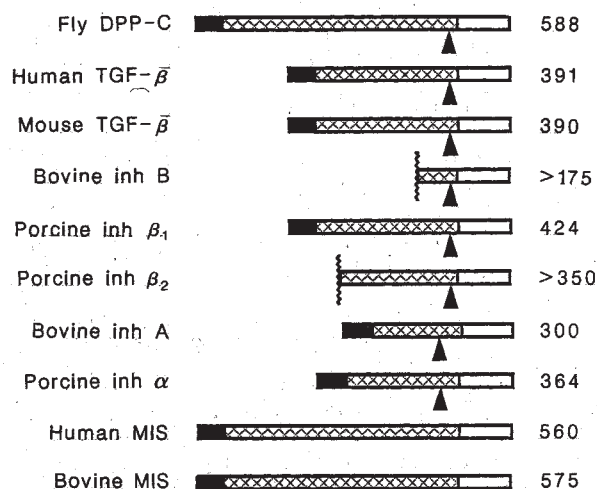


Fig. 3 Schematic diagram of the structure of the related genes of the TGF- β gene family. The heterodimeric form of inhibin consists of A and B chains (cows⁶) or α and β chains (pigs⁷). The β gene has duplicated to produce β_1 and β_2 which show 70% amino-acid homology. The B gene in cows most resembles the β_1 gene; an analogue of β_2 has not yet been identified in cows. Open box, the highly related C-terminal ends of these genes. Triangle, known or potential cleavage sites^{41,42}, which release the secreted polypeptide (MIS members are not thought to be cleaved). Solid box at the 5' end, the putative signal sequence found in each gene. The cDNAs of the porcine and bovine inhibin are truncated in the ORF. Numbers indicate length in amino acids.

indicating that the composite cDNA contains all but a few hundred base pairs of the complete transcript. The level of this transcript peaks at several times in development (early embryos, early third instar larvae, early pupae and adults)³. Its temporal and spatial expression partially overlaps with the expression patterns of other DPP-C mRNAs.

The composite cDNA has an open reading frame (ORF) of ~1,800 base pairs (bp) surrounded by ~1,200 bp of 5' and ~1,000 bp of 3' untranslated sequences. It is probably complete at its 3' end as it has a terminal poly-A tract which is preceded by a consensus polyadenylation signal. A termination codon, which places a limit on the 5' extent of the ORF, is located at position 1,170. The two AUG codons at the 5' end of the ORF are the only reasonable candidates for the translation initiation codon. The first, at nucleotide position 1,188, would produce a protein 588 amino acids long (with a relative molecular mass (M_r) of 66,000 and a pI of 10.8) whereas the second, at position 1,587, would produce a protein 455 amino acids long (with M_r 52,000 and a pI of 10.6). The sequences surrounding these AUG codons conform to the general eukaryotic initiation codon consensus sequence CC[G/A]CCATGG¹³ and to the specific *Drosophila* consensus sequence ANN[C/A]A[A/C]-[A/C]ATGN (D. Cavener, personal communication). Beginning with the second AUG of the ORF, the codon usage is very strongly biased towards preferred *Drosophila* codons¹⁴, but there is no significant codon bias in the interval between these two possible initiation codons using this particular codon bias table. However, the sequence following the first AUG of the ORF, but not the second, matches the hydrophobic characteristics of an amino-terminal signal peptide^{15,16}. The ORF contains several potential sites for glycosylation and dibasic proteolytic processing.

Protein homologies were sought in the Dayhoff protein data base (National Biomedical Research Foundation Protein Sequence Data Base of the Protein Identification Resource, Washington, DC) using the Lipman-Pearson protein sequence similarity program¹⁷. No homologies to known *Drosophila* genes

were identified. Further examination of the DPP-C protein reveals no regions enriched for any single amino acid^{18,19} (such as the *opa* sequence²⁰), no homoeobox homology^{21,22} and no homology to the finger motif first identified in *Xenopus* TFIIIA²³, and which has been found in the *Krüppel*²⁴ and *serendipity*^{25,26} loci of *Drosophila*.

The protein database search and subsequent literature searches did reveal homology between the sequence of the carboxy-terminus (~100 amino acids) of the DPP-C ORF and the comparable regions of a family of four related mammalian polypeptides (Fig. 3): TGF- β (refs 4, 5), inhibin-A and inhibin-B (refs 6, 7) and MIS (ref. 10). Whereas MIS appears not to be cleaved, the other proteins are derived from precursors which, after proteolysis at Arg-Arg residue pairs near their C-termini, produce secreted biologically active polypeptides. The homologies of these proteins with each other and with the deduced DPP-C polypeptide are restricted to these C-terminal regions (Fig. 4). The DPP-C polypeptide contains three sets of Arg-Arg residue pairs in the vicinity of the expected cleavage position, based on the sites of cleavage of TGF- β , inhibin-A and inhibin-B. We propose that the DPP-C protein, like its mammalian relatives, exists as a dimer because all seven cysteines that occur after the cluster of potential DPP-C cleavage sites are conserved in all the related proteins. Note that the three highly conserved domains surround these seven cysteines (Fig. 4).

It is not clear if the only biological activity of the DPP-C polypeptide resides in the putative C-terminal fragment. There is high conservation of the N-terminal third of the TGF- β precursors of mice and men⁵, even though there is no corresponding homology of this region to other members of the gene family. Derynck *et al.*⁵ suggest that important functions may reside within this region of TGF- β .

Thus, we speculate that the translated product of the 4.5-kb DPP-C mRNA will begin at the first AUG of the ORF and will produce a 588 amino-acid precursor which has an N-terminal secretion signal sequence. Following processing and secretion, a C-terminal polypeptide (~100 amino acids long) will be produced that contributes one or both subunits to a biologically active dimeric protein. If further parallels with TGF- β and inhibin exist, then the biologically active DPP-C polypeptide would affect cell growth and patterning by binding to a cell surface receptor²⁷.

The proposed structure of the DPP-C product supports the possibility that cells in a specific region of an imaginal disk (the *dpp* focus) are the source of a DPP-C polypeptide that promotes growth of surrounding target cells. In mosaic animals containing both wild-type cells and *dpp* mutant cells (of a genotype precluding wing blade development), the normal morphogenesis of the wing blade depends on having *dpp*⁺ cells occupying the *dpp* focus of the developing wing imaginal disk (ref. 28 and L. Posakony, personal communication). When the focus contains *dpp*⁺ cells, large patches of genetically *dpp*⁻ cells are present in a normal wing blade.

Our developmental studies of the DPP-C clearly imply that it has an important role in the growth regulation of *Drosophila*, including determination of dorsal versus ventral in the early embryo, and correct morphogenesis of the imaginal disks in the larva. Likewise, all the related mammalian proteins are involved in growth regulation and/or differentiation. TGF- β , a homodimer, can inhibit or stimulate the growth of cells or cause them to differentiate²⁹⁻³¹. Inhibin, found in ovarian fluid, inhibits pituitary follicle-stimulating hormone (FSH) release and is a heterodimer of two distantly related polypeptides, both being homologous to TGF- β ⁶⁻⁹. Recently, porcine inhibin β chain dimers have been shown to have the opposite effect to inhibin $\alpha\beta$ heterodimers, stimulating secretion of FSH^{8,9}. MIS is expressed in mammalian testes, causing the regression of the Müllerian ducts during the early development of the male reproductive system^{10,11}. It is not clear how to relate these

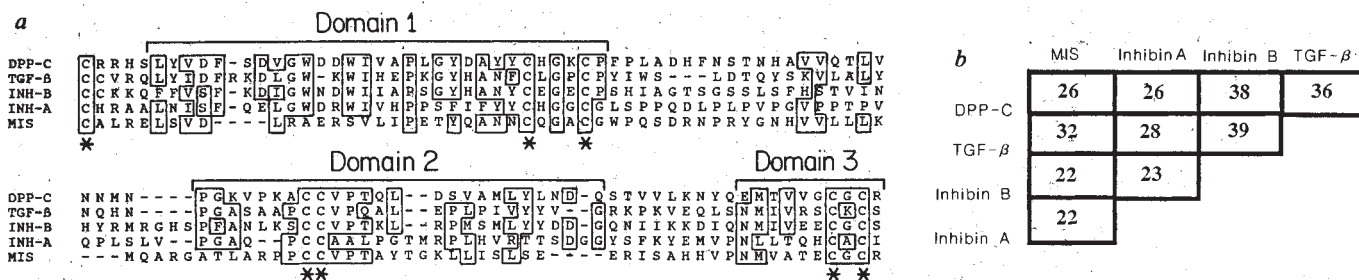


Fig. 4 The homologies between the related genes. **a**, Amino-acid sequence alignment of the C-terminal ends of the polypeptides, beginning with the first conserved cysteine and ending at the translation terminator. A single representative gene from each family of the related class was chosen for alignment (TGF- β from human⁴, inhibin A and B from cow⁷ and MIS from human¹⁰). Note that similar numbers of deletions in each case were invoked in aligning the related genes. Positions where at least three of the five aligned sequences agree are enclosed in boxed areas. Below each conserved cysteine is an asterisk. TGF- β and inhibin B possess a conserved cysteine not shared by other members at position 2. There are three polypeptide domains that exhibit high sequence conservation (brackets). The regions outside these domains are essentially unconserved. **b**, Table indicating the percentage amino-acid homology among the genes. Pairwise comparisons were made beginning with the first conserved cysteine residue. Percentages were calculated by scoring the number of identities between two sequences disregarding positions in which one or both sequences had deletions added to improve the alignment.

activities in any detail; the gene family was identified on the basis of the protein sequence homology searches following cDNA sequence analysis, rather than on physiological similarities.

The identification of the homologies between the DPP-C and the mammalian TGF- β gene family demonstrates that an ancestral gene must have been present before arthropods and vertebrates diverged. It is striking that the degree of homology between C-termini of DPP-C and either TGF- β or inhibin-B is as strong as the homology between these latter two proteins. Furthermore, the three are more closely related to each other than they are to inhibin-A or MIS. This suggests that the gene family had at least two members at the time of arthropod-vertebrate divergence. Given the recent discovery of the TGF- β gene family, including the present report on the DPP-C representative, the discovery of additional family members may be expected. It will be intriguing to see if functions analogous to TGF- β , inhibin or MIS will be identifiable in *Drosophila* and, if a DPP-C-like protein will be found in vertebrates. Because of the involvement of the DPP-C in the fundamental and evolutionarily ancient process of embryonic dorsal-ventral determination, we suggest that a vertebrate DPP-C analogue might exist.

We thank L. Kauvar for providing the *Drosophila* embryonic cDNA library from which E3 and E55 were isolated and J. Ostell, R. Lewontin, C. A. Hutchison III and the Molecular Biology Computer Research Resource for providing sequence analysis software. We thank members of this laboratory for encouragement and specifically Ron Blackman, Maggie de Cuevas, Holly Irick, Macy Koehler, Leila Posakony and Laurel Raftery for critical comments regarding the manuscript. Supported by a NIH grant GM 28669. R.W.P. was the recipient of an NIH postdoctoral fellowship.

Received 29 August; accepted 6 October 1986.

- Spencer, F. A., Hoffman, F. M. & Gelbart, W. M. *Cell* **28**, 451-461 (1982).
- Segal, D. & Gelbart, W. M. *Genetics* **109**, 119-143 (1985).
- Gelbart, W. M. *et al. Cold Spring Harb. Symp. quant. biol.* **50**, 119-125 (1985).
- Derynck, R. *et al. Nature* **316**, 701-705 (1985).
- Derynck, R., Jarrett, J. A., Chen, E. Y. & Goeddel, D. V. *J. biol. Chem.* **261**, 4377-4379 (1986).
- Mason, A. J. *et al. Nature* **318**, 659-663 (1985).
- Forage, R. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3091-3095 (1986).
- Vale, W. *et al. Nature* **321**, 776-779 (1986).
- Ling, N. *et al. Nature* **321**, 779-782 (1986).
- Cate, R. L. *et al. Cell* **45**, 685-698 (1986).
- Budzik, G. P., Powell, S. M., Kamagata, S. & Donahoe, P. K. *Cell* **34**, 307-314 (1983).
- Poole, S., Kauvar, L. M., Drees, C. & Kornberg, T. *Cell* **40**, 37-43 (1985).
- Kozak, M. *Nucleic Acids Res.* **12**, 857-872 (1984).
- Pustell, J. DNA sequence analysis programs marketed by International Biotechnologies, Inc., New Haven, Connecticut (1986).
- Perlman, D. & Halvorson, H. O. *J. molec. Biol.* **167**, 391-409 (1983).
- Watson, M. E. *Nucleic Acids Res.* **12**, 5145-5164 (1984).
- Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
- Schneuwly, S., Kuroiwa, A., Baumgartner, P. & Gehring, W. *EMBO J.* **5**, 733-739 (1986).

- Laughon, A. & Scott, M. P. *Nature* **310**, 25-31 (1984).
- Wharton, K. A., Johansen, K. M., Xu, T. & Artavanis-Tsakonas, S. *Cell* **43**, 567-581 (1985).
- McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa & Gehring, W. J. *Nature* **308**, 428-433 (1984).
- Scott, M. P. & Weiner, A. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4115-4119 (1984).
- Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609-1614 (1985).
- Rosenberg, U. B. *et al. Nature* **319**, 336-339 (1986).
- Vincent, A., O'Connell, P., Gray, M. R. & Rosbash, M. *EMBO J.* **3**, 1003-1013 (1984).
- Vincent, A., Colot, H. V. & Rosbash, M. *J. molec. Biol.* **186**, 149-166 (1985).
- Sporn, M. B., Roberts, A. B., Wakefield, L. M. & Assoian, R. K. *Science* **233**, 532-534 (1986).
- Spencer, F. A. Ph.D. thesis, Harvard Univ. (1984).
- Roberts, A. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 119-123 (1985).
- Masui, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 2438-2442 (1986).
- Lawrence, D. A. *Biol. Cell* **53**, 93-98 (1985).
- Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211-214 (1976).
- Wagh, P. V. & Bahl, O. P. *CRC crit. Rev. Biochem.* **10**, 307-377 (1981).
- Bankier, A. T. & Barrell, B. G. in *Techniques in Nucleic Acid Biochemistry* B5 (ed. Flavell, R. A.) 1-34 (Elsevier, Limerick, 1983).
- Dretzen, G., Bellard, M., Sassone-Corci, P. & Chambon, P. *Analyt. Biochem.* **112**, 295-298 (1981).
- Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).
- Hutchison III, C. A. *Nucleic Acids Res.* **14**, 1917 (1986).
- Staden, R. *Nucleic Acids Res.* **10**, 4731-4751 (1982).
- Staden, R. *Nucleic Acids Res.* **12**, 499-503 (1984).
- Mills, D. R. & Kramer, F. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2232-2235 (1979).
- Noda, M. *et al. Nature* **295**, 202-206 (1982).
- Amara, S. G., Jonas, V. & Rosenfeld, M. G., Ong, E. S. & Evans, R. M. *Nature* **298**, 240-244 (1982).

Two variant surface glycoproteins of *Trypanosoma Brucei* of different sequence classes have similar 6 Å resolution X-ray structures

P. Metcalf, M. Blum, D. Freymann, M. Turner* & D. C. Wiley

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

* Merck, Sharp and Dohme Research Laboratories, PO Box 2000, Rahway, New Jersey 07065, USA

Antigenic variation in the African trypanosome is mediated through changes in the composition of the surface coat^{1,2}. By controlling expression of the major surface protein, the variant surface glycoprotein (VSG), from a repertoire of perhaps 1,000 different genes³ the organisms exhibit a series of antigenically distinct coats and evade the host's immune system. We have determined the 6 Å resolution structure of a *T. brucei* variant surface glycoprotein, ILTat 1.24, using X-ray crystallography. The crystallized protein consists of the N-terminal two-thirds of the intact VSG which has a relative molecular mass (M_r) of 60,000

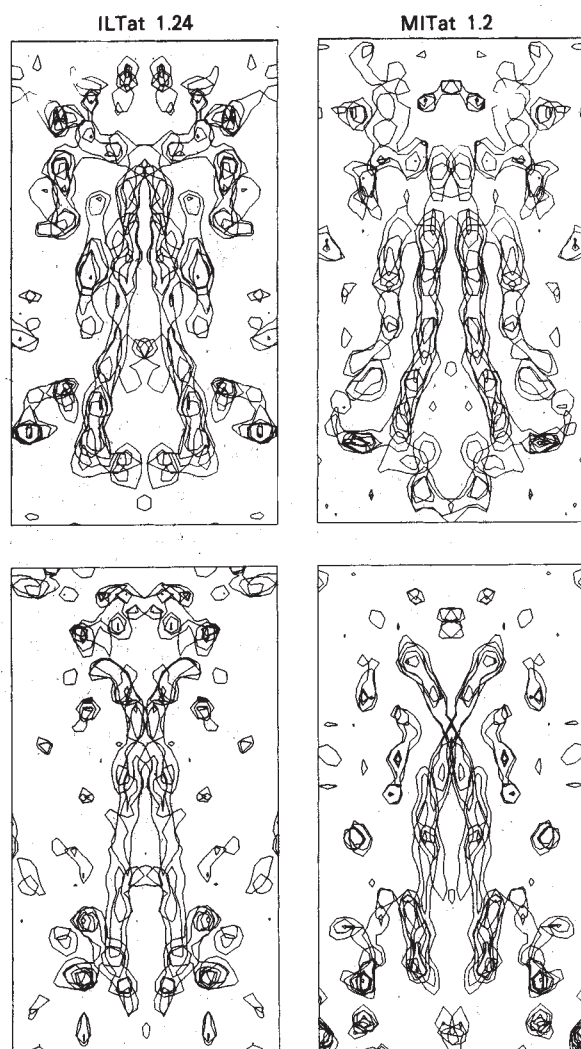


Fig. 1 Central lateral sections showing the α -helical core of the molecules: Left ILTat 1.24, right MITat 1.2. Each diagram ($110 \text{ \AA} \times 60 \text{ \AA}$) consists of 5 superimposed $2 \text{ \AA}$ sections of the $6 \text{ \AA}$ resolution electron density map contoured at 1 and 2 standard deviations above the mean density. Noise in low density regions of the electron density map was reduced using an interactive solvent flattening algorithm²¹. The relative orientation of the twofold averaged ILTat 1.24 map with respect to the MITat 1.2 map was determined by maximizing the cross-correlation of their electron densities (see Table 1).

(60K). The structure, which includes a $90 \text{ \AA}$ long α -helical bundle, is strikingly similar to that of the N-terminal fragment of VSG MITat 1.2 (ref. 4). Although most known VSG sequences show little similarity of primary sequence in the N-terminal domain, the similarity between the structure of a Class I (ILTat 1.24) and a Class II (MITat 1.2) VSG antigen suggests that VSGs may share a common tertiary structure.

VSG molecules are attached to the cell surface via a C-terminal glycolipid anchor⁵⁻⁷. The protein can be purified in a soluble form (sVSG), lacking the lipid, and is a dimer in solution^{8,9}. Proteolysis experiments suggest the existence of two domains¹⁰. Significant sequence homologies between VSGs are confined to the C-terminal domain and are found largely in the last 50 amino acids^{11,12}. Based on these homologies and two different patterns of cysteine residues in the C-terminal domain VSGs have been divided into two classes having C-terminal aspartic acid (Class I) or serine (Class II)^{11,12}. The N-terminal, variable domain comprises two-thirds of the molecule. No amino-acid sequence

Table 1 Summary of $6 \text{ \AA}$ data from ILTat 1.24 VSG crystals

	Native	Hg
Cell dimensions ($\text{\AA}$)	55.1, 98.8, 172.9	55.1, 98.6, 172.4
No. of reflections collected	11452	11577
Symmetry R factor	0.037	0.055
No. of unique reflections in R factor	2394	2472
No. of reflections phased		2470
Mean figure of merit		0.68
Kraut R		11.8
$\langle F_h \rangle / \langle E \rangle$		1.98
$CC_{\text{two-fold}}$		0.28
CC_{cross}		0.27

$$\text{Symmetry R factor} = \frac{\sum_{hkl} \sum_{\text{obs}} (I_{\text{obs}} - \langle I_{hkl} \rangle)}{\sum_{hkl} \sum_{\text{obs}} I_{\text{obs}}}$$

$$\text{Kraut R} = \frac{\sum [|F_{\text{ph}}(\text{obs}) - F_{\text{ph}}(\text{calc})|]}{\sum [F_{\text{ph}}(\text{obs})]} \times 100,$$

where F_{ph} is the derivative structure factor;

$$\langle F_h \rangle / \langle E \rangle = \frac{\text{r.m.s. heavy-atom structure factor}}{\text{r.m.s. lack of closure error}};$$

$$CC = \frac{\langle \rho_1 \rho_2 \rangle - N \langle \rho_1 \rangle \langle \rho_2 \rangle}{[(\langle \rho_1^2 \rangle - N \langle \rho_1 \rangle^2)(\langle \rho_2^2 \rangle - N \langle \rho_2 \rangle^2)]^{1/2}}$$

where for $CC_{\text{two-fold}}$, ρ_1 , ρ_2 are electron densities at positions related by rotation of 180° around the dimer axis; for CC_{cross} , ρ_1 = ILTat 1.24 electron density, ρ_2 = MITat 1.2 electron density. The space group is $P2_12_12_1$; unit cell, $a = 55.1$, $b = 98.9$, $c = 172.9 \text{ \AA}$. Bipyramidal crystals grow to $0.4 \times 0.3 \times 0.3 \text{ mm}^3$ and diffract to $2.5 \text{ \AA}$. The asymmetric unit of the crystal contains one copy of the N-terminal dimer. An isomorphous mercury derivative was obtained by soaking the crystals for 6 h in 65% ammonium sulphate, 5 mM MES pH 6.5, 0.2% azide and 0.1 mM K_2HgI_4 . The $4.2 \text{ \AA}$ data were collected on a multiwire area detector (Nicolet/Xentronics, Madison, WI) mounted 12 cm from the crystal, using a rotating anode X-ray source (Elliot GX-13) with 100 mm focus and Franks double mirror optics^{17,18}. Each data set consists of 1,500 data frames (0.083 degree oscillations of the crystal exposed for 90 s) which were collected from crystals at 4°C over a 50 h period and processed on-line¹⁸. The diffraction data were divided into groups consisting of integrated spot intensities from 50 consecutive frames and these groups were scaled to each other¹⁹. Harker sections from the isomorphous and anomalous difference Patterson maps were interpreted as being consistent with a single Hg site at fractional unit cell coordinates 0.076, 0.190, 0.082. The heavy atom parameters were refined²⁰ and single isomorphous replacement and anomalous scattering phases calculated for reflections of resolution less than $6 \text{ \AA}$. Correlation coefficients were calculated inside a cylinder with radius of $30 \text{ \AA}$ and length of $110 \text{ \AA}$ oriented along the dimer axis.

homologies larger than two residues and no significant homology at the DNA level are evident in the first 350 amino acids of the N-terminal domains¹². Nevertheless, three observations argue for an underlying conservation of structure: (1) the apparent conservation of four cysteines in several Class I VSGs¹, (2) the existence of sevenfold periodicities in hydrophobicity indicating stretches of interlocking α helices in four sequences¹³, and (3) the fact that the first 40 amino acids of 18 N-terminal sequences have significant similarities, if conservative amino-acid substitutions are allowed¹⁴. The lack of agreement in secondary structure predictions for the amino-terminal domain of 9 sequences¹⁵ has, however, been interpreted to suggest multiple folding patterns.

Sequence analysis of a partial complementary DNA clone of ILTat 1.24 showed that it belonged to the Class I family (M. Carrington, P. Liberator and M.T., unpublished). The complete sequence of the MITat 1.2 VSG has been deduced by sequence analysis of a cDNA clone, establishing it as a member of Class

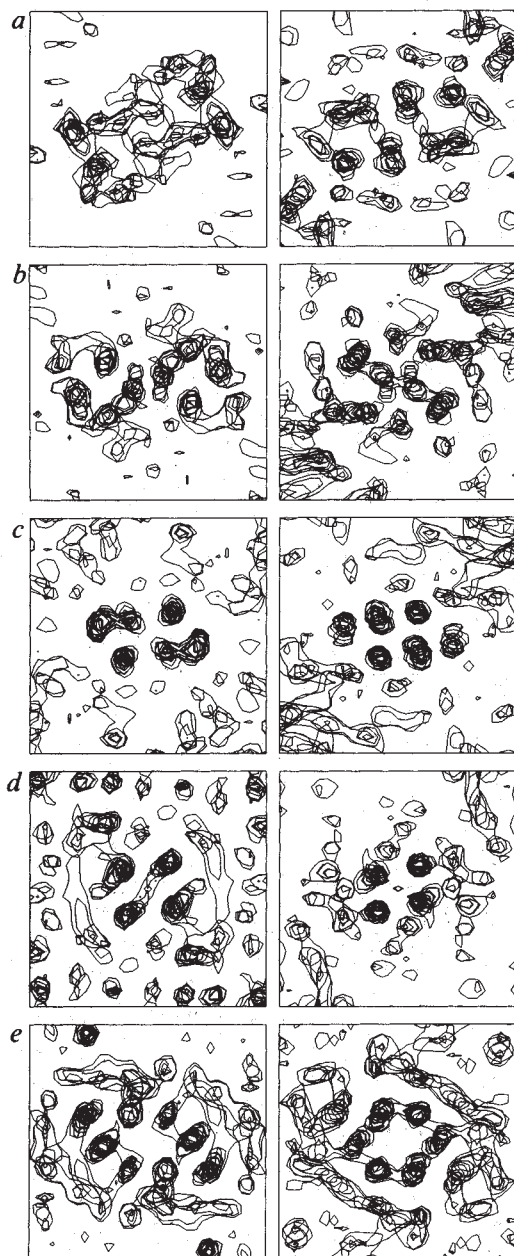


Fig. 2 Axial sections: Left ILTat 1.24, right MITat 1.2 *a*, 80–96 Å; *b*, 64–80 Å; *c*, 48–64 Å; *d*, 32–48 Å; *e*, 16–32 Å from the bottom of the region shown in Fig. 1. Each diagram (60 Å × 60 Å) consists of 8 superimposed 2 Å sections each contoured at 1 and 2 standard deviations above the mean density.

II (M. Carrington, I. Roditi and M.T., unpublished). ILTat 1.24 sVSG was prepared as described¹⁶. Purified sVSG was crystallized by vapour diffusion of a 40 mg ml⁻¹ solution (50 mM MES, pH 6.5, 0.2% azide, 30% (NH₄)₂SO₄) against 60%-saturated ammonium sulphate in the same buffer. Protein recovered from the ILTat 1.24 crystals had a substantially lower *M_r* (46K) than the native purified protein (60K). The amino-terminal sequence of the crystalline fragment was found to be the same as the intact sVSG (T-FGVK).

The initial 6 Å electron density map (crystallographic data are summarized in Table 1) showed strong features suggesting a rod-shaped molecule packed with its long axis at ~40° from the *c* axis in the *b*, *c* plane. The electron density sectioned perpendicular to this molecular axis showed features recognizable as the sixfold and fourfold helical bundle regions seen in

the low resolution structure of MITat 1.2 (ref. 4). Comparison of the map to the current model of the MITat 1.2 VSG structure (D.F. *et al.*, unpublished) allowed a more accurate determination of the dimer's non-crystallographic twofold axis. Orientation of this axis was refined by maximizing the twofold electron density correlation in a cylindrical region of the map and the density was then averaged about the refined axis.

The MITat 1.2 and ILTat 1.24 N-terminal dimers are both elongated molecules with a length of ~100 Å and a maximum width of 60 Å. The twofold axes of the dimers are parallel to the long axes of the molecules. The dominant features of both maps (Figs 1, 2) are long rods of high electron density, which are characteristic of α -helices at this resolution. The arrangements of these helices, which form the central core of the molecules, are similar: two helices extend nearly 90 Å, forming a fourfold bundle around the dimer axis in the middle third of the molecule (Fig. 2*c*, *d*). A third helix is packed against the long helices to form a sixfold bundle in the bottom third of the molecule (Fig. 2*e*). A fourth helix is packed against the top of the fourfold bundle (Fig. 2*b*, *c*). These features account for at least 40% of the polypeptide chain.

The top of the central helical core is capped by a region containing what appears to be non- α -helical density in both molecules. This cap extends ~20 Å and, viewed from above, has a rectangular cross-section of ~30 Å × 50 Å (Fig. 2*a*). At this resolution we cannot interpret the cap region density or establish the connectivity of the α -helices.

The extent of structural homology between the two VSGs at 6 Å suggests that there may be an underlying structural motif within the VSGs. The primary selection pressure for the survival of VSG sequences may be their ability to conform to the constraints of this structure^{22,23}. This similarity also leaves open the question of what significance, if any, the classification of VSGs, based on C-terminal homologies, has to the structure of the N-termini. Fragments of four other VSGs have been crystallized²⁴ and it is likely that at least some of these will eventually yield further structures. Should the sequences of ILTat 1.24 and MITat 1.2 prove to be as dissimilar as expected from our current knowledge of VSG sequences, the similarity of their structures suggests the possibility of creating a useful structural database from the large number, possibly 1,000, of disparate VSG sequences that may correspond to a single protein fold.

We thank Judith Creighton for technical assistance in the preparation of VSGs and Anne Gurnett and Michael Silver for their interest and help. D.C.W. acknowledges support from NIH (AI-21324-02) and NSF (CHE-8509574). This work was initiated by the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases (project No. 810278 and No. 800328).

Received 23 July; accepted 17 October 1986.

1. Cross, G. A. M. *Phil. Trans. R. Soc. Lond.* B307, 3–12 (1984).
2. Turner, M. J. *Phil. Trans. R. Soc. Lond.* B307, 27–40 (1984).
3. Van der Ploeg, L. H. T. *et al. Nucleic Acids Res.* 10, 5905–5923 (1982).
4. Freymann, D., Metcalf, P., Turner, M. J. & Wiley, D. *Nature* 311, 167–169 (1984).
5. Cardoso de Almeida, M. L. & Turner, M. J. *Nature* 302, 349–352 (1983).
6. Ferguson, M. A. J. & Cross, G. A. M. *J. biol. Chem.* 259, 3011–3015 (1984).
7. Ferguson, M. A. J., Haldar, K. & Cross, G. A. M. *J. biol. Chem.* 260, 4963–4968 (1985).
8. Auffret, C. A. & Turner, M. J. *Biochem. J.* 193, 647–650 (1981).
9. Gurnett, A. M., Raper, J. & Turner, M. J. *Molec. biochem. Parasitol.* 18, 141–153 (1986).
10. Johnson, J. G. & Cross, G. A. M. *Biochem. J.* 178, 689–697 (1979).
11. Holder, A. A. & Cross, G. A. M. *Molec. biochem. Parasitol.* 2, 135–150 (1981).
12. Rice-Ficht, A. C., Chen, K. K. & Donelson, J. E. *Nature* 294, 53–57 (1981).
13. Cohen, C. *et al. Nature* 311, 169–171 (1984).
14. Olafson, R. W. *et al. Molec. biochem. Parasitol.* 12, 287–298 (1984).
15. Lalor, T. M. *et al. Proc. natn. Acad. Sci. U.S.A.* 81, 998–1002 (1984).
16. Cross, G. A. M., *J. Cell Biochem.* 24, 79–90 (1983).
17. Durbin, R. M. *et al. Science* 232, 1127–1132 (1986).
18. Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. *J. appl. Cryst.* (in the press).
19. Fox, G. C. & Holmes, K. C., *Acta cryst.* 20, 886–891 (1966).
20. Dickerson, R. E., Weinzierl, J. E. & Palmer, R. A. *Acta cryst.* B24, 997–1003 (1968).
21. Wang, B. C. *Meth. Enzym.* 115, 90–112 (1985).
22. Turner, M. J., *Biochem. Soc. Symp.* 49, 169–181 (1984).
23. Turner, M. J., *Annls Inst. Pasteur, Paris (Immunol.)* 136C, 41–49 (1985).
24. Metcalf, P. *et al.* (manuscript in preparation).

The colour of Sirius in the sixth century

ON the basis of Gregory of Tours' description of the star called 'Rubeola' or 'Robeola' in his *'De cursu stellarum ratio'*¹⁻⁴, Schlosser and Bergmann⁵ have proposed that Sirius appeared red in the sixth century, an appearance which would raise serious astrophysical questions. From close examination of the text, the identification of Sirius with 'Rubeola' cannot be supported.

Gregory's treatise appears in eight manuscripts, two of which^{1,2} include the astronomical portions, and two printed editions^{3,4}, both of which include identifications of the constellations by the astronomer, J. F. Galle. Galle identified 'Rubeola' as Arcturus, a nearby companion as η Boo, and the central star of Gregory's constellation 'Quinio' (the five) as 'undoubtedly Sirius'. He considered 'Rubeola' to be Arcturus because of its colour (which we should not take into account) and its monthly appearances as recounted by Gregory.

There are several independent sources of evidence in *'De cursu stellarum ratio'*. In the first section Gregory describes a number of constellations, giving their month-by-month visibility from heliacal rising to heliacal setting, depicting (in the eighth-century Bamberg MS) the constellations as they appear rising on the eastern horizon, and in some cases expressing their declination in terms of the path that the Sun takes in a given month.

In the second section Gregory uses these constellations to time the pre-dawn office of matins^{6,7} for each month, telling the reader the number of psalms to be chanted, beginning when a star rises or more commonly when it reaches a given hour of the day. The latter expression

apparently indicates the time when the star reaches the position that the Sun would attain at the given hour of the day, that is, so many hours after the star rises. These two sections employ general terms rather than exact measurements and are not fully consistent; it therefore seems wisest to consider all the available evidence rather than a selected set.

From the fact that 'Rubeola' is twice described as rising heliacally in September and is also visible for an hour after sunset the same evening, we can compute limits for right ascension and declination. From the contradictory data for a maximum visibility of 8 hours from December through February, which I would question on textual grounds, alternative limits can be computed (W. Schlosser and W. Bergmann, personal communication).

The Local Sidereal Time of rising (LST_r) can be computed from 'Rubeola's rising one month before 'Stefadium' (Corona Borealis), from the duration of visibility each month between its rising and sunrise, and from the number of psalms that can be sung after its rising or after its appearance at a given place in the sky. The Local Sidereal Time of setting (LST_s) can be determined in like manner from the duration of visibility each month between sunset and the star's setting. Positions were adjusted for precession and proper motion; times of rising and setting were computed for the latitude of Tours (47.3° N) without corrections for refraction. First nighttime visibility is taken as when a bright star is on the horizon⁸.

Turning to the constellation 'Quinio', which has also been identified as Sirius³, we can determine its declination because it follows the path that the Sun does in February. Because it rises one hour after a pair of stars that can be identified on the basis of their declination, their rising

time, and their depicted orientation to the horizon as α and β Canis minoris, we can determine 'Quinio's' LST_r. Lastly, we must not ignore the depiction of 'Quinio', which matches the pattern of Sirius and its companions at their rising^{1,4}.

Comparing the limits thus obtained with the positions of the stars computed for AD 600 at the latitude of Tours (Table 1). It is apparent that 'Quinio' is Sirius and 'Rubeola' is most likely Arcturus, a star widely noted for its ruddy colour. The importance of Arcturus for monastic timekeeping is reinforced by its appearance as one of seven bright stars in an eleventh-century monastic timetable from northern France, in which Sirius does not appear⁹. Gregory's description of 'Rubeola' cannot be added to the ancient reports of a reddish Sirius¹⁰.

STEPHEN C. MCCLUSKEY

*Program in the History
of Science and Technology,
West Virginia University,
Morgantown,
West Virginia 26506, USA*

1. Bamberg, Staatsbibliothek, *patres* 61 (HJ.IV.15), ff. 75v-82v.
2. Vatican City, Biblioteca Apostolica Vaticana, ms Urbinas lat. 67, ff. 120v-126v.
3. Gregorii Turonensis, *'De Cursu Stellarum Ratio'* (ed. Haase, F.) (Max, Vratislaviae, 1853).
4. Gregorii Turonensis, *Opera*, pars II, *Miracula et Opera Minora* (eds Arndt, W. & Krusch, B.) (MGH, Scr. rer. merov., I), (Hahn, Hanover, 1885).
5. Schlosser, W. & Bergmann, W. *Nature* **318**, 45-46 (1985).
6. de Vogüé, A. (ed. transl.) *La Règle du Maître* Vol. 2, 168-187 (du Cerf, Paris, 1964).
7. Fry, T. (ed. transl.) *The Rule of Saint Benedict* 202-205 (Liturgical Press, Collegeville, Minnesota, 1981).
8. Bruin, F. *Proc. ned. Akad. Wet.* **B82**, 385-410 (1979).
9. Poole, R. J. *theol. Stud.* **16**, 98-104 (1914).
10. Brecher, K. in *Astronomy of the Ancients* (eds Brecher, K. & Feitrag, M.) 95-102 (MIT Press, Cambridge, 1979).

SCHLOSSER and Bergmann¹ (hereafter SB) have recently presented new evidence apparently corroborating the well-known Greek-Roman and Babylonian references to a red-coloured Sirius in antiquity^{2,3}. The purpose of this letter is to show that they have erroneously confounded Sirius with Arcturus (α Boo), a star which is indeed reddish ($B - V = 1.2$). SB based their analysis on *'De cursu stellarum ratio'* (hereafter CS), a text written sometime between 575 and 582 by Georgius Florentius Gregorius, Bishop of Tours. This text was published twice in the previous century by Haase⁴ and Krusch⁵ (hereafter H and K). The text (CS 20, 36-39) refers to a *'stella splendida'*, named 'Robeola' (=reddish). According to the editions of H and K, 'Robeola' = Arcturus, but SB (without even acknowledging the earlier work of HK) argue rather that 'Robeola' = Sirius.

Their novel identification is partly based on the singular description of 'Robeola' as a *'stella splendida'*, which, as SB argue, can apply only to the brightest star in the

Table 1 Positions of 'Rubeola' and 'Quinio'

From <i>'De cursu stellarum ratio'</i>	'Rubeola'	Sirius	Arcturus
September heliacal rising			
Dec. (deg.)	>14.3	-15.75	26.95*
RA (h:min)	10:47-12:32	5:43	13:12
December opposition (doubtful)			
Dec. (deg.)	<0	-15.75*	26.95
RA (h:min)	4:40-10:50	5:43*	13:12
Appears one month before 'Stefadium'			
LST _r (h:min)	0:56-4:51	0:55	4:58
Monthly visibilities			
LST _r (h:min)	3:18-7:16	0:55	4:58*
LST _s (h:min)	12:02-19:07	10:32	21:26
Times for Nocturnal Psalms			
LST _r (h:min)	2:19-4:57	0:55	4:58
	'Quinio'		
Dec. (deg.)	-7.0 to -16.6	-15.75*	26.95
LST _r (h:min)	23:57-1:22	0:55*	4:58

* Star within the range computed for the constellation.

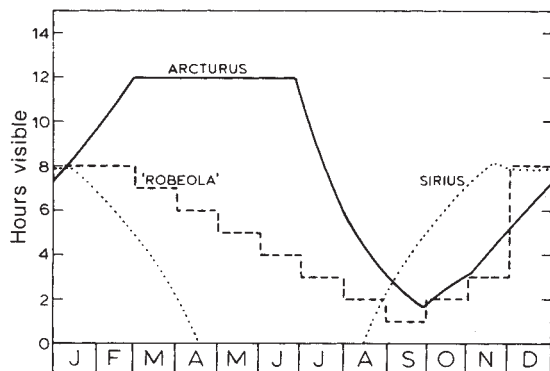


Fig. 1 Visibility times of 'Robeola' as given by Gregory (broken line; the missing value for May is interpolated from the adjacent months) and those theoretically expected for Arcturus (solid line) and Sirius (dotted line) at Tours ($\phi = 47.37^\circ$). Stellar coordinates adopted for the epoch 500¹⁵. Unequal-hour scheme adopted. Visibility limits adopted were: altitude Sun $< -12^\circ$ and altitude star $> +5^\circ$.

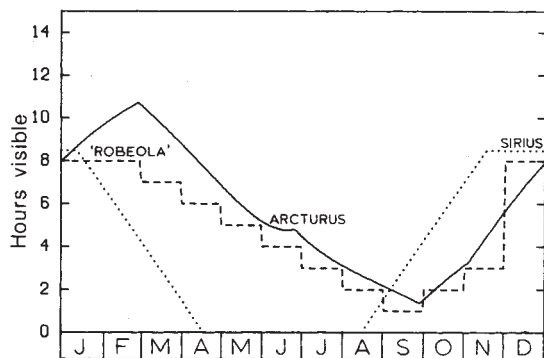


Fig. 2 As for Fig. 1 for the equal-hour scheme.

sky (Sirius) and not to any other fainter star. But although Sirius features prominently in the skylore of many ancient cultures (such as Egypt, Persia, Greece and Rome), this is not a general rule. In other, more or less contemporary cultures (Mesopotamia, India and China), Sirius played a much less important role and was never mentioned as a specially bright or conspicuous star. Thus Gregory's reference to Robeola as a '*stella splendida*' does not automatically implicate Sirius.

The second argument of SB for rejecting the Arcturus identification is based on the times of visibility of 'Robeola' throughout the year, which, as SB state, only fits Sirius. Figures 1 and 2 show the visibility times of 'Robeola' as given by Gregory (CS 20) compared with that theoretically expected for Arcturus and Sirius at the latitude of Tours. In Fig. 1, hours of unequal length or seasonal hours (12 in each night, irrespective of the season) are assumed and it

shows that Arcturus and Sirius both fit very badly. Figure 2, where hours of equal length are assumed, shows that Arcturus now almost perfectly fits the data given by Gregory, while Sirius still disagrees completely. Regardless of the hour-scheme adopted (a study of the visibility data of the other constellations listed by Gregory in fact shows that the latter scheme is correct), Sirius always remains invisible each year from May to July, whereas 'Robeola' should be visible for at least 3 hours each night during the same period. This crucial and very elementary fact is nowhere raised nor discussed by SB.

Gregory further states (CS 20) that 'Robeola' is visible twice each night in September; shortly after sunset as well as shortly before sunrise. Figure 3 shows that again Arcturus fits perfectly (as its heliacal rising in late September falls before its heliacal setting in early November)⁶, whereas Sirius fails miserably (heliacal ris-

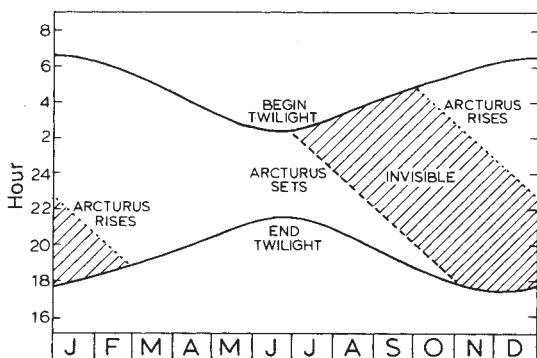


Fig. 3 Rising and setting times for Arcturus at Tours in mean local time (hours of equal length). Visibility criteria as in Figs 1 and 2. Note that from late September until early November, Arcturus is visible shortly after sunset as well as shortly before sunrise.

ing in early August falls after its heliacal setting in late April)⁶. Finally, it may be noted that 'Robeola' as Arcturus fits more logically in the sequence of constellations listed by Gregory, namely Bootes ('Robeola', CS 20), Corona Borealis ('symma', CS 21), Lyra ('omega', CS 22), Cygnus ('crux maior', CS 23), Delphinus ('crux minor', CS 24), Aquila ('trion', CS 25), Pisces(?) ('signum Christi', CS 26), Hydra ('anguis', CS 27), Pleiades ('Pliadas', CS 28), Gemini ('arte feretrum', CS 29), Leo ('falx', CS 30), Virgo (CS 31), Scorpius(?) ('quinio', CS 32) and finally Ursa Major ('plaustrum', CS 33). Except for the last constellation (probably added later), there is a systematic sequence from the northern to the southern constellations. HK identified Sirius as one of the five stars ('quinio', CS 32) circling in the southern skies, but I find no evidence that Sirius is mentioned anywhere in the CS.

It is therefore evident that SB have failed to prove that 'Robeola' = Sirius. There are no grounds for questioning the conventional identification with Arcturus, proposed long ago by HK. It is also clear that Gregory's text presents no new evidence for a red-coloured Sirius in antiquity. Regarding the other often-quoted classical references for a red colour, other contemporary texts (occidental as well as oriental) are known describing Sirius as white or bluish⁷⁻⁹. The reddish colour occasionally attributed to Sirius was probably caused by atmospheric extinction when Sirius was observed close above the horizon at its heliacal rising (around 19 July). This important celestial event was attentively watched for each year in the ancient world, as the apparent colour and brightness of Sirius was believed to presage the coming year's weather, yield of the crops, political fortune of the ruler and health of his people¹⁰. From a study of Hellenistic and Babylonian data on the first and last visibility of planets and bright stars, Vogt¹¹ and Schoch¹²⁻¹⁴ demonstrated that the apparent brightness of Sirius then did not differ significantly from its present value ($V = -1.5$) and that it was about as bright as Jupiter when near solar conjunction ($V \approx -1.5$ to -1.9), but certainly not as bright as Venus ($V \approx -4.2$), as SB state. There is at present no evidence that any significant change took place in the apparent colour or brightness of Sirius during the past few millennia.

R. H. VAN GENT

Sonnenborgh Observatory,
Zonnenburg 2,
3512 NL Utrecht,
The Netherlands

- Schlosser, W. & Bergmann, W. *Nature* 318, 45-46 (1985).
- Schiaparelli, G. V. *Atti R. Accad. Sci. Lett. Arti Agiate (III)* 2 (2) (1896); 3 (1) (1897).
- See, T. J. *J. Astr. Nachr.* 229, 245-272 (1927).

4. Haase, F. S. *Georgii Florentii Gregorii Turonensis episcopi liber ineditus de cursu stellarum ratio qualiter ad officium implendum debeat observari sive de cursibus ecclesiasticis* (Ios, Max. et Soc., Bratislava, 1853).
5. Krusch, B. *Monumenta Germaniae Historica, Scriptores Rerum Merovingicarum*, Vol. I, Pt ii, 404-422 (Hahn, Hanover, 1885).
6. Bickerman, E. J. *Chronology of the Ancient World* 2nd edn, 112 (Cornell University Press, Ithaca, 1980).
7. van Gent, R. H. *Nature* 312, 302 (1984).
8. Tong, B. T. *Nature* 319, 532 (1986).
9. van Gent, R. H. *Zenit* 13, 212 (1986).
10. Gundel, W. in *Paulys Real-encyclopädie der classischen Altertumswissenschaft - Neue Bearbeitung*, Vol. IIIA, Cols 314-351 (Metzlersche, Stuttgart, 1925).
11. Vogt, H. *Sber. heidelb. Akad. Wiss., Philos.-hist. Kl.*, Abh. 15 (1920).
12. Schoch, K. *Mon. Not. R. astr. Soc.* 84, 731-734 (1924).
13. Schoch, K. *Arcus Visionis in Babylonian Observations, with Tables of the Babylonian Calendar* (University Observatory, Oxford, 1924).
14. Schoch, K. *Planeten-tafeln für Jedermann*, sp. xlv (Linser, Berlin, 1927).
15. Baehr, U. *Veröff. Astron. Rechen-Inst. Heidelberg* no. 3 (1955).

SCHLOSSER AND BERGMANN REPLY—McClusky and van Gent advance the same arguments that led Krusch and Galle a century ago to the assignation 'Robeola' = Arcturus. Since then we have become much better informed about the intentions Gregory pursued in his astronomical writings. We have to distinguish between two different levels, A and B.

A. In most of his writings, Gregory proved to be an enthusiastic 'amateur astronomer', as we would put it today. He made an immense number of observations (from solar eclipses to polar lights) and described them in his *Historia Francorum* and the *Seventh Wonder of the World*, the latter being the basis of the Arcturus hypothesis. As has been shown by Bergmann and Schlosser¹, Gregory generally reported correctly what he saw, but not necessarily the exact temporal circumstances.

So it is not surprising to find that the fine drawings of constellations in the Bamberg codex are not paralleled by numerical data of the same quality. This is demonstrated by van Gent's Fig. 2 and McClusky's Table 1 (upper part), which are anything but proof for the Arcturus hypothesis. Although the slope of the 'Robeola' data in Fig. 2 corresponds somewhat better to Sirius, we never have based our identification on data like these, but on Gregory's level B (see below).

In this connection, we flatly reject van Gent's 'convincing' Fig. 1 as a non-historic discussion of mediaeval manuscripts. Gregory used (as the middle age generally) 'horae inaequales', i.e. an 'unequal hour scheme'. The night was always divided into twelve hours, even in December, without respect to its physical length.

Krusch and Galle (and today McCluskey) assigned Sirius and some stars around to Gregory's constellation 'quino'. Not much is said about this constellation, but Gregory tells us that these stars were rising three hours after 'falx', which had its heliacal rise in early August. Since at that time Sirius had (for the latitude of Tours) its first morning-rise too, 'quino' was rising

several hours after Sirius and an identification of Sirius with 'quino' seems improbable.

These examples—and more could be given—simply demonstrate, that the *Seventh Wonder of the World* alone is not sufficient to allow a unambiguous identification of Merovingian sky with modern constellations.

B. In the first place, Gregory was a bishop rather than an 'amateur astronomer'. In AD 567 the Council of Tours took place. His predecessor, Euphronius, presided over this meeting. The minutes of this Council define in chapter 18 the number and length of the pre-dawn prayer sequences to be conducted in the monasteries. Gregory's '*De cursu stellarum ratio qualiter ad officium implendum debeat observari*' (about 'On the Course of Stars and how to use it for Divine Services') is the practical regulation of this decree. It should be kept in mind that the monks had no alarm-clocks at their disposal and were punished 'on bread and water' between one day and one week for violating this decree.

Obviously, these prayer instructions were based on observations and, indeed, no inconsistencies show up in this list arranged in monthly order. The standard phrase for 'Robeola' (several psalms to be chanted plus following matin) is repeated for other stars giving the following durations:

Vega	January	4.3 h
Vega	February	6.0 h
Capella	April	1.7 h
Capella	May	2.7 h
Capella	June	3.9 h
Hyades	August	3.6 h

All times are for $\phi = 47^\circ$, AD 580, Sun's depression 12° , elevation of stars 7° . For Capella, which Gregory correctly notes as circumpolar, lower culmination is used. Furthermore, all dates refer to the fifth of each month to allow for a small 'jitter' in Gregory's calendar.

Crucial are the September and October apparitions of Rubeola. Whereas Sirius = 'Robeola' fits nicely into the scheme

Sirius	September	1.7 h
Sirius	October	4.4 h

Arcturus = 'Robeola' does not:

Arcturus	September	rising $\frac{1}{2}$ h after the Sun
Arcturus	October	10 min.

We do not consider the 'red Sirius problem' finally solved by our paper. But we do think that our evaluation of this interesting manuscript took into account every foreseeable aspect of history and

astronomy, and hope that it might stimulate further work in this field.

W. SCHLOSSER
W. BERGMANN

*Astronomisches Institut,
Ruhr-Universität Bochum,
Postfach 102148,
D-4630 Bochum 1, FRG*

I. Bergmann, W. & Schlosser, W. *FRANCIA—Forschungen zur westeuropäischen Geschichte* 13 (Herausgegeben vom Deutschen Historischen Institut, Paris, 1986).

Early forest clearance and the environment in south-west Uganda

THE pollen diagram from Ahakageyezi Swamp, Uganda can be interpreted in a radically different manner from that of Hamilton, Taylor and Vogel¹. The stable isotope readings from the two lowest dated levels clearly indicate that the swamp vegetation at those times (4,670 years BP and 3,520 years BP) was dominated by C_4 plants, most probably savanna type grasses². At some point before 2,800 yr BP, the stable isotope results indicate a change to a more 'normal' C_3 type vegetation. These data suggest that below ~6-7 metres, drier climatic conditions allowed savanna type grasses to dominate the vegetation on the swamp surface. However, the stratigraphy, loss-on-ignition, charcoal and pollen analytical results provide a more detailed picture of environmental change.

The stratigraphy records two clay bands, between 10 m and 9.5 m and between 8.25 m and 7.5 m, which are reflected in matching peaks in inorganic matter content. On East African mires such bands have been found to occur during drier conditions^{3,4}. The peaks of 'microscopic charcoal' in these bands would tend to suggest the natural burning of the vegetation during drier periods.

The pollen diagram in the early period shows no pollen that could be regarded as a positive indicator of anthropogenic disturbance or agriculture. The interval 10-9.5 m, which has high levels of inorganic matter and charcoal, is dominated by Gramineae pollen, reflecting the dominance of grasses on the swamp surface and in the catchment. The only other individual taxon to reach a significant percentage is *Alchornea*. This could be *A. cordifolia* or *A. hirtella*, which have different ecologies. *A. cordifolia* is found today in secondary forest, swamp forest and forest-edge situations⁵⁻⁷, whereas *A. hirtella* is an understory tree of moist forest^{5,7}. The site at Ahakageyezi Swamp is above the known, present-day altitudinal limit for *A. cordifolia*⁸, but with a different climate, altitudinal limits would also change. A further complicating factor is that *Alchornea* is a moderately well dispersed pollen type and therefore need not necessarily be growing at or near the site

to produce the pollen counts found in the lowest levels of the profile⁹. Pollen Groups 4 and 5 also make a significant showing, but since these groups are composed of a mixture of species found in swamp and secondary forest and of ubiquitous species, their indicator value is unclear⁵⁻¹¹.

At 9.35 m, inorganic matter content and charcoal decrease. This change is echoed in the pollen diagram by increases in *Myrica*, Myrtaceae, *Schefflera* and ferns, all of which are commonly found in swamp forest^{5,7,10}, as well *Macaranga* and *Olea*, which are common taxa of secondary or scrub forest⁵⁻⁷. *Alchornea* rises to reach its maximum level in the profile at 8.75 m. Pollen Groups 4 and 5 increase but their indicator value is unclear. Gramineae show a significant decrease. These factors suggest that wetter conditions allowed swamp forest to develop on the valley floor and forest on the slopes.

At 8.25 m, the inorganic matter content and charcoal again increase, the swamp forest species decline and Gramineae increase, indicating a return to drier conditions. This change occurs at ~3,700 years BP, which is widely regarded as marking a major change to drier conditions in East Africa^{4,12}.

The dramatic rise in swamp forest species at 7.5 m (~3,400 yrs BP) indicates a return to wetter conditions.

It is clear from these data that the early part of this pollen diagram reflects the influence of climatic change and not anthropogenic influence as the authors suggest. Further stable isotope analyses from this core should prove informative.

R. ALAN PERROTT

Tropical Palaeoenvironments
Research Group,
School of Geography,
University of Oxford,
Oxford OX1 3TB, UK

1. Hamilton, A., Taylor, D. & Vogel, J. C. *Nature* **320**, 164-167 (1986).
2. Winter, K., Troughton, J. H. & Card, K. A. *Oecologia* (Berlin) **25**, 115-123 (1976).
3. Perrott, R. A. *7th Biennial Conf. AMQUA 153*, (University of Washington, Seattle, 1982).
4. Hamilton, A. C. *Environmental History of East Africa* (Academic, London, 1982).
5. Hamilton, A. *Uganda Forest Trees* (Makerere, Uganda, 1981).
6. White, F. *The Vegetation of Africa* (UNESCO, Paris, 1983).
7. Lind, E. M. & Morrison, M. E. S. *East African Vegetation* (Longman, London, 1974).
8. Hamilton, A. thesis, Makerere University, Uganda (1970).
9. Hamilton, A. *Palaeocol. Afr.* **7**, 45-149 (1972).
10. Dale, I. R. & Greenway, P. J. *Kenya Trees and Shrubs* (Buchanan's Kenya Estates Limited, Nairobi, 1961).
11. Agnew, A. D. Q. *Upland Kenya Wild Flowers* (Oxford U. Press, 1974).
12. Perrott, R. A. *Palaeocol. Afr.* **14**, 77-83 (1982).

TAYLOR HAMILTON REPLY—Perrott's criticism of our interpretation of the work from Ahakageyezi¹ rests on the assertion that it was two early dry phases which were responsible for the changes in vegetation, not the destructive influence of man.

These 'more arid' periods are marked by abundant grass pollen and have given

$\delta^{13}\text{C}$ values for two sediment samples of -18.2 and -15.9%, contrasting with $\delta^{13}\text{C}$ values of -27.0, -26.0 and -26.9% for three samples higher up the core. Perrott assumes that C_4 plants, in particular savanna grasses, are responsible for the less negative $\delta^{13}\text{C}$ values. A swamp surface, like that at Ahakageyezi (which has one of the highest rates of peat accumulation known from Africa²), is hardly a suitable substrate for savanna type grasses; it cannot be compared to seasonally inundated floodplains where C_4 grasses can grow³ and which have inorganic soils⁴. Measurements of $\delta^{13}\text{C}$ are available for some C_4 grasses: *Andropogon amethystinus* (-10.9%), *Pennisetum clandestinum* (-11.8%), *Sporobolus africanus* (-11.0 to -11.3%) and *Themeda triandra* (-11.4%) (D. Harkness, personal communication); all of which are more positive than our stable isotope readings said to represent savanna grasses.

The $\delta^{13}\text{C}$ values from Ahakageyezi can be compared with those from core samples from Kamiranzovu Swamp, Rwanda⁵. The upper three samples are from sediment rich in fossil wood and yield $\delta^{13}\text{C}$ values close to those obtained from similar sediments from Kamiranzovu (-26.9, -27.0%). The lower two samples contain abundant herbaceous fossils and have similar $\delta^{13}\text{C}$ values to those of sediments rich in herbaceous fossils and Cupraceae pollen at Kamiranzovu (=16.6, -17.9%). Sedge pollen is one of the most common types in the lower samples from Ahakageyezi, but it is less frequent than at Kamiranzovu and we envisage a wetter environment. Under such moist conditions the algal contribution to the sediment is likely to have been great. Typically planktonic algae have $\delta^{13}\text{C}$ values between -12.0 and -23.0% (ref. 6).

Aridification, beginning ~3,700 BP, has been reported from many sites in Africa⁵. Levels of inorganic material appear to be generally less since 3,700 BP than they are before it, the latter being a period marked by a more moist climate. Inorganic inwash as a consequence of forest clearance has been previously detected at one other site in southwest Uganda⁷ and is well known in other parts of the world. Sedimentation in East African basins is complex²; we contend that at Ahakageyezi sediments rich in inorganic material have accumulated when the climate was wetter or following human disturbance.

The absence of pollen definitely attributable to cultivated plants is a feature common to most pollen diagrams from tropical Africa^{5,8}. Many possible food crops at Ahakageyezi are either poor pollen producers (for example, legumes, root crops) or have pollen indistinguishable by light microscopy from that of wild species (e.g. millets, sorghum). Reliance is thus placed on agrestal and secondary forest taxa, along with the declines exhibited by many

moist lower montane forest pollen spectra. Contrary to Perrott's opinion, decreases even in such uncommon pollen types as Ebenaceae, *Musanga/Myrianthus* and *Parinari* are significant given their under-representation in the pollen rain and in view of the pollen diagram's absolute nature.

We still believe that the evidence, as a whole, signifies early forest destruction and burning and is a manifestation of human activity. We agree that there is no direct evidence as to why this might have occurred, although cultivation is most likely. We are continuing with our own research in the area, and await confirmation from archaeologists.

D. TAYLOR
A. HAMILTON

*Department of Environmental Studies,
University of Ulster, Coleraine,
Co. Londonderry BT52 1SA, UK*

1. Hamilton, A., Taylor, D. & Vogel, J. C. *Nature* **320**, 164-167 (1986).
2. Hamilton, A. & Taylor, D. in *Sedimentation in the African Rifts* (eds Frostick, L. E. et al.) *Geol. Soc. Spec. Publ.* **23** (in the press).
3. Thompson, K. & Hamilton, A. C. in *Mires: Swamp, Bog, Fen and Moor*, B. *Regional Studies* (ed. Gore, A. J.) 331-373 (Elsevier, Amsterdam, 1983).
4. Lind, E. M. & Morrison, M. E. S. *The Vegetation of East Africa* (Longman, London, 1974).
5. Hamilton, A. C. *Environmental History of East Africa* (Academic, London, 1982).
6. Smith, B. N. & Epstein, S. *Plant. Physiol.* **47**, 380-384 (1971).
7. Morrison, M. E. S. & Hamilton, A. C. *J. Ecol.* **62**, 1q-31 (1974).
8. Kendall, R. L. *Ecol. Monogr.* **39**, 121-176 (1969).
9. Hamilton, A. C. *Palaeocol. Afr.* **7**, 45-149 (1972).

Copies of articles
from this publication
are now available
from the UMI Article
Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnType ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
Title _____
Institution/Company _____
Department _____
Address _____
City _____ State _____ Zip _____
Phone (_____) _____

**UMI Article
Clearinghouse**

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Products to watch for in the new year

Convenience is the watchword for this first week of 1987, with everything from gels to cells.

ARTIFICIAL intelligence (AI) technology is springing up in many laboratories with the advent of several new programs for specific biological applications. One of the newest is StrateGene, a **genetic engineering computer workstation** developed by IntelliGenetics, currently released as a beta-test version (Reader Service No. 100). StrateGene runs on the Xerox 1100 series of computer workstations, and is based on KEE, an AI tool that assists in the creation of expert systems. The system's knowledge includes reasoning about experimental methods in molecular biology, and its screen display combines graphics of whole plasmid maps along with selected regions of their nucleotide sequences. IntelliGenetics sees the \$50,000–60,000 (US) system as a useful tool for both keeping track of recombinant plasmids and for designing new ones.

A new automated system developed by Oncor, Inc. **streamlines the electrophoresis and transfer elements** of the Southern procedure (Reader Service No. 101). The



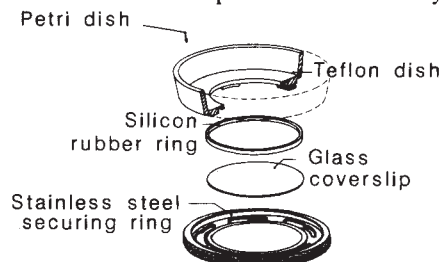
Cuts time required for Southern hybridization from days to hours.

\$5,700 (US) Probe Tech 1 processes up to two gels of eleven lanes each within four hours, according to Oncor. Oncor says the increase in speed does not sacrifice accuracy, though, and that the Probe Tech 1 detects DNA fragments in the range of 100 to 25,000 base pairs, capturing 90 per cent on the nylon membrane.

The Immunoassays Simulation Program developed by Clin-Lab Services Ltd, sets out to **take some of the trial and error out of immunoassays** (Reader Service No. 102). The program uses the law of mass action to simulate standard curves, and can compare the results of up to three different assay systems at once. The results are then displayed in either a graph or table format, with up to 28 variations of parameters, along with the precision profile and corresponding error envelope. Response–error relationships may be plotted and interrelated with the standard curve simulations to evaluate the implica-

tions of differing assay designs on performance. The £350 (UK) program runs on an IBM-PC, and can be tried out with a free demonstration diskette, available from Clin-Lab Services upon request.

High-magnification **microscopy of growing tissue cultures** is possible with Medical Systems Corporation's Teflon culture dish developed at the University



For watching cultures as they grow.

of Leiden, East Germany (Reader Service No. 103). The \$295 (US) dish consists of a Teflon ring that holds a glass coverslip on which the culture is grown, and is designed for use with Medical Systems' LUCB-1 Leiden tissue culture bath. The glass of the coverslip makes it suitable for $\times 100$ oil immersion, phase-contrast and $\times 400$ interference-contrast microscopy.

Clonetics Corporation has introduced what may be the first commercially available normal **human epidermal cell culture** (Reader Service No. 104). Epipack was created to provide researchers with a more realistic model of human skin than can be achieved with the abnormal human or animal cell lines currently used for the purpose. The \$250 (US) keratinocyte culture is provided in a serum-free medium



Better than abnormal cells.

and shipped in a T-25 flask, together with 500 ml of growth factor-supplemented medium and premixed formulations for subculturing. Clonetics claims Epipack provides a 100 million cell count culture after incubating 10 days, and that seven passages are possible before senescence, for a total of 20 to 25 cell doublings.

However, Clonetics does not recommend using the cells for assays beyond the third passage.

To **reduce the neurotoxin hazard** involved in handling powdered acrylamide, National Diagnostics offers AcrylaGel, a 30 per cent acrylamide solution in distilled, deionized water (Reader Service No. 105). The solution is stabilized by a gaseous inhibitor to eliminate self-polymerization, that, says National Diagnostics, is completely removed upon deaeration of the gel casting solution. A one-litre bottle, containing the equivalent of 300 g of acrylamide, costs £24 (UK). National Diagnostics claims its AcrylaGel solution is purer than powdered acrylamide, containing zero acrylic acid content, and no trace metals.

More and more products for AIDS research are entering the market, and DuPont Ltd is keeping pace with other producers. Its latest AIDS product is a **nick translation DNA probe for detecting HTLV-III** (Reader Service No. 106). The probe identifies the presence of HTLV-III by hybridization with viral DNA, with an observed sensitivity of 10^4 to 10^5 copies of positive control DNA in an overnight exposure, according to DuPont. The £169 (UK) kit comes with HTLV-III purified DNA fragment, standard curve positive control DNA, a NENSORB-20 cartridge, control plasmid DNA, and the necessary labelled nucleotide solutions and restriction endonucleases. Ten separate ^{32}P nick translations can be done with this system, and 120 to 450 samples can be screened per kit.

ADVERTISEMENT

LIPOSOMES

REPRODUCIBLE – HOMOGENEOUS –
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIABORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.47

BIOTEC 87

International Congress with Exhibition Bioengineering and Genetechnology

○ Including the **BIOTEC Specialised Fair** covering all sectors of biotechnology. Now presented on a larger surface area by far more exhibitors.

○ With **BIOTEC Congress** for an up-to-date dialogue between scientists and economists.

Lectures on Biotechnology for Environmental Protection by:

Dipl.-Ing. W. A. Enger, Dr. S. P. Etheridge, Prof. Dr. Th. Leisinger, Dr. V. Melin, Dr. G. Müller, Prof. Dr. K. N. Timmis, R. B. Volesky, Dr. J. T. Wilson, Prof. Dr. H. Sahm.

Lectures on the subject of Biosensors by:

Dr. B. Danielsson, R. J. Erickson, Prof. Dr. G. G. Guilbault, Prof. Dr. I. J. Higgins, Dr. E. Hall (Mrs.), Prof. Dr.-Ing. I. Karube, Dr. Charles R. Keese, Dr. M. Mason.

If you want to keep pace with developments on one of the fastest growing markets of our time, **BIOTEC 87** should take absolute priority over all other appointments – for visitors and exhibitors alike.

Secure your share of success with us!

You are welcome to **BIOTEC 87** held March 17 to 19, 1987.

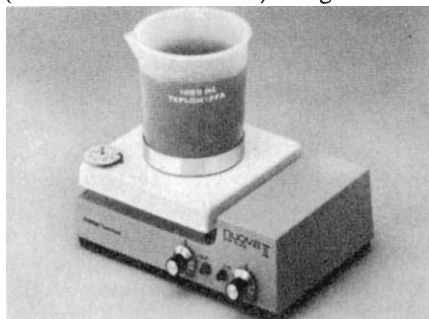


NOWEA
DUSSELDORFER MESSEN

Düsseldorfer Messgesellschaft mbH - NOWEA -
P.O. Box 32 02 03 · D-4000 Düsseldorf 30
Phone (0211) 45 60-01 · Telex 8 584 853 mes d

Reader Service No.4

The pitfalls associated with heating glass beakers can be avoided with Nalge's new Type HP **heatable plastic beaker** (Reader Service No. 107). Nalge's beaker



Doesn't crack up under the heat.

is composed of Teflon-PFA with a graphite base fused to the bottom, that ensures uniform heating while eliminating the problems of bumping, cracking, and splashing. Nalge says the Teflon composition also makes its beaker ideal for ultra-pure applications, because it does not absorb contaminants and is not susceptible to scratches that can induce precipitation. The beakers are claimed to achieve evaporation to dryness twice as fast as their glass counterparts. The beakers come in 100 ml, 250 ml, 400 ml and 1,000 ml sizes, and cost \$70–175 (US).

Rainin Instrument Company (RIC) has announced new **hexagonal columnar hydroxyapatite (HCA) columns** for purification of monoclonal antibodies using HPLC (Reader Service No. 108). Rainin says HCA is an ideal column material for this application because of its stability over a wide pH range and its high loading



An old material gets a new look.

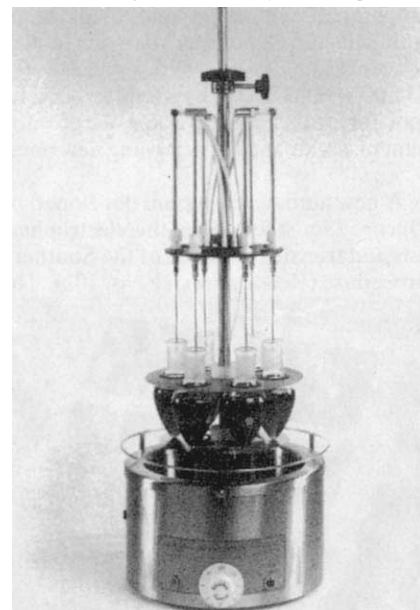
capacity. The columns are available in stainless steel or glass, and have flow rates of up to 1.5 ml min⁻¹ per square centimetre. Sold in a wide variety of sizes, the costs of the columns range from \$600 to \$6,500 (US).

Bioscan, Inc.'s new Quick-Count QC-2000 **benchtop counter** allows researchers to have their biological samples and count them too (Reader Service No. 109). No sample preparation is involved, so samples may be counted in working vials or tubes without transfer or dilution. The counter provides an innovation over conventional liquid scintillation counting, according to Bioscan, because the QC-2000 performs counts in minutes rather than hours, and is easily incorporated into

a lab routine. The \$2,350 (US) unit is microprocessor-based, and provides readings in c.p.m. or d.p.m. using ³²P or ¹²⁵I isotopes.

Post-expressional protein modification is a key step in the preparation of biotechnological products. To meet these needs, Fluka has recently released a free comprehensive catalogue of 700 protein modification reagents (Reader Service No. 110). A sampling of Fluka's chemicals include reagents for cross-linking, group-specific modification, protein iodination and photoaffinity labelling.

For the fast and pure **evaporation of aqueous samples**, Organomation Associates, Inc. has introduced the Aqua-Vap analytical water evaporator (Reader Service No. 111). The \$960 (US) evaporator



Dries up samples in no time.

uses six volumetric or pear-shaped 25–100 ml flasks, and is designed to evaporate 60 ml of water per hour. The machine functions by blowing nitrogen or air onto the surface of the sample to be evaporated, while the flask is immersed in a steam bath. The flow of the gas onto the sample can be controlled by a needle valve over each flask, to regulate the rate of evaporation.

Setting it straight

The BB Form Petri dish imprinter reviewed in the 2 October issue of *Nature* (323, 473; 1986) was developed by C.R. Bader and D. Bertrand of Centre Medical Universitaire in Geneva in collaboration with Mecanex S.A. It is distributed by World Precision Instruments. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.